

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 31, 2006, 08:42:29 ; Search time 10148 Seconds
(without alignments)
10920.465 Million cell updates/sec

Title: US-10-759-813-1
Perfect score: 1733
Sequence: 1 gcataaaggaaatggagc.....tcttttcaaatccgaaaaa 1733

Scoring table: IDENTITY_NUC
Gapop 10_0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues
Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_env.*
- 2: gb_pat.*
- 3: gb_ph.*
- 4: gb_pl.*
- 5: gb_pr.*
- 6: gb_ro.*
- 7: gb_sts.*
- 8: gb_sy.*
- 9: gb_un.*
- 10: gb_vi.*
- 11: gb_ov.*
- 12: gb_hcg.*
- 13: gb_in.*
- 14: gb_om.*
- 15: gb_ba.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1733	100.0	1733	2	AX366075 Sequence
2	1729.8	99.8	1758	4	AF406732 Euphorbia
3	410	23.7	1538	2	CS223906 Sequence
4	400.4	23.1	1673	2	CS223962 Sequence
5	400.4	23.1	1673	2	CS223964 Sequence
6	400.4	23.1	1673	2	CS223984 Sequence
7	399.4	23.0	1781	2	E63735 Cytochrome
8	399.4	23.0	1781	2	AR205821 Sequence
9	399.4	23.0	1781	4	AF122821 Capsicum
10	398.8	23.0	1673	2	CS223966 Sequence
11	393.2	22.7	1610	2	CS223950 Sequence
12	393.2	22.7	1610	2	CS224014 Sequence
13	388.8	22.4	1566	2	CS223944 Sequence
14	388.8	22.4	1576	2	CS223902 Sequence
15	388.8	22.4	1581	2	CS223904 Sequence
16	388.8	22.4	1664	2	CS223952 Sequence
17	388.6	22.4	1657	2	CS223948 Sequence
18	387.6	22.4	1701	4	STU296346 Solanum t

19	387.4	22.4	1789	4	DQ335785	DQ335785 Medicago
20	384.8	22.2	1754	4	GM450CP3	Y10490 G. max mRNA
21	384.2	22.2	1559	2	CS224040	CS224040 Sequence
22	376.8	21.7	1509	4	AB028462	AB028462 Petunia x
23	376.6	21.7	1784	2	CS249182	CS249182 Sequence
24	369.8	21.3	1665	2	A29004	A29004 pCGP147 DNA
25	369.8	21.3	1665	2	I47318	I47318 Sequence 29
26	367.4	21.2	1666	2	AR030921	AR030921 Sequence
27	367.4	21.2	1666	2	I28313	I28313 Sequence 24
28	366.6	21.2	1800	4	GM450CP7	Y10493 G. max mRNA
29	364.6	21.0	1641	4	AF000403	AF000403 Lotus jap
30	361	20.8	1506	4	AY596975	AY596975 Centaureia
31	357.6	20.6	1572	2	CS223930	CS223930 Sequence
32	355.6	20.5	1515	4	AF368376	AF368376 Nicotiana
33	351.8	20.3	1682	4	CR0238612	CR0238612 Catharan
34	350.4	20.2	1641	4	SCU48434	U48434 Solanum cha
35	347.4	20.0	1599	2	CS223912	CS223912 Sequence
36	344.2	19.9	1593	2	CS223910	CS223910 Sequence
37	344.2	19.9	1694	2	CS224048	CS224048 Sequence
38	339.6	19.6	1823	2	CS223960	CS223960 Sequence
39	338	19.5	1767	2	CS223956	CS223956 Sequence
40	336.4	19.4	1736	2	CS223958	CS223958 Sequence
c	41	315	18.2	132419	4	CR954196 Medicago
42	308	17.8	1691	4	AF022459	AF022459 Glycine m
43	296.8	17.1	1564	2	CS223914	CS223914 Sequence
44	295.2	17.0	1560	2	CS223916	CS223916 Sequence
45	294.6	17.0	1541	4	AY192573	AY192573 Catharan

ALIGNMENTS

RESULT 1
AX366075
LOCUS AX366075
DEFINITION Sequence 1 from Patent WO208269.
ACCESSION AX366075
VERSION AX366075.1 GI:18697502
KEYWORDS Euphorbia lagascae
SOURCE Euphorbia lagascae
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Euphorbiaceae; Euphorbioideae;
Euphorbiae; Euphorbia.

REFERENCE 1
AUTHORS Cahoon,E.B.
TITLE A cytochrome p450 enzyme associated with the synthesis of _g(d)712
-epoxy groups in fatty acids of plants
JOURNAL Patent: WO 0208269-A 1 31-JAN-2002;
E. I. du Pont de Nemours and Company (US)
FEATURES
source Location/Qualifiers
1..1733
/organism="Euphorbia lagascae"
/mol_type="unassigned DNA"
/db_xref="taxon:54672"

ORIGIN
Query Match 100.0%; Score 1733; DB 2; Length 1733;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1733; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GCATTAAGGAAATGGAGCAGAAAATCTCTTTCCGAGCAATTTAATAAGTTTCT	60
Db	1	GCATTAAGGAAATGGAGCAGAAAATCTCTTTCCGAGCAATTTAATAAGTTTCT	60
QY	61	GCTTGTGTTTAATCTTAGTAGTAGTCATGAGGTGTGGAGAAACAGAAATCCACCTCCAGG	120
Db	61	GCTTGTGTTTAATCTTAGTAGTAGTCATGAGGTGTGGAGAAACAGAAATCCACCTCCAGG	120
QY	121	GCCATGGAAGTTTCTTATCATAGGTAATCTTCCTCATTTTATCTCATCTTCTGATCTAGG	180
Db	121	GCCATGGAAGTTTCTTATCATAGGTAATCTTCCTCATTTTATCTCATCTTCTGATCTAGG	180

Qy	181	CCATGAACGCTTTTAGAGCCTTGGCTCAAAATTTATGGACCTGTATGAGCTCTTCAAAATGG	240
Db	181	CCATGAACGCTTTTAGAGCCTTGGCTCAAAATTTATGGACCTGTATGAGCTCTTCAAAATGG	240
Qy	241	CCAAGTTTCAGCTGTGTGTCATTTCTTACGCTGAAGCAGCCAAAGAGGTTTGAACACTCA	300
Db	241	CCAAGTTTCAGCTGTGTGTCATTTCTTACGCTGAAGCAGCCAAAGAGGTTTGAACACTCA	300
Qy	301	GGCTGATGCCCTTCGCCCAACGCCCTATCGTCTTGACGCGACAGATGCTGTTTTATATCG	360
Db	301	GGCTGATGCCCTTCGCCCAACGCCCTATCGTCTTGACGCGACAGATGCTGTTTTATATCG	360
Qy	361	GAAGATGCTCTTGTGTTTCATATGGAGATCACTGGAGGAGATGAAGAAAATTTGGAT	420
Db	361	GAAGATGCTCTTGTGTTTCATATGGAGATCACTGGAGGAGATGAAGAAAATTTGGAT	420
Qy	421	ACTTGAATTTCTGAGTGCCCAAAAAGTTCAATCTTCCAGGTTAATCCGAGGAGAGAAAT	480
Db	421	ACTTGAATTTCTGAGTGCCCAAAAAGTTCAATCTTCCAGGTTAATCCGAGGAGAGAAAT	480
Qy	481	GGAGATGCCATCACATTCCTCCGTTCGAAAGCCGGATCTCCGGTCAATATTTACAAAGAT	540
Db	481	GGAGATGCCATCACATTCCTCCGTTCGAAAGCCGGATCTCCGGTCAATATTTACAAAGAT	540
Qy	541	CATTATATGGCATTAATTTTCGATCATGATGAAGAACATCCGTTGGTAAATTTGAAGCAAAA	600
Db	541	CATTATATGGCATTAATTTTCGATCATGATGAAGAACATCCGTTGGTAAATTTGAAGCAAAA	600
Qy	601	AGAAAGATTGCTGAGTGTTCGGATGCGATCAATGAGCAGCGAGGTTTGGCACC GC	660
Db	601	AGAAAGATTGCTGAGTGTTCGGATGCGATCAATGAGCAGCGAGGTTTGGCACC GC	660
Qy	661	AGACGCTTTTCCGAGCTGGAATTAATCTCACTATATCAATGGAGCTGAGTCAAAACCCAG	720
Db	661	AGACGCTTTTCCGAGCTGGAATTAATCTCACTATATCAATGGAGCTGAGTCAAAACCCAG	720
Qy	721	CGGTTTGATCAGGAGATTGACGATATATCTTGAAGAGATTTCTTAATGAACACAAAGCCAA	780
Db	721	CGGTTTGATCAGGAGATTGACGATATATCTTGAAGAGATTTCTTAATGAACACAAAGCCAA	780
Qy	781	TAAGCTTTTGAAGCGGATACTTAATGGATGTTCTATTGATCTTCAAAAATGGAAA	840
Db	781	TAAGCTTTTGAAGCGGATACTTAATGGATGTTCTATTGATCTTCAAAAATGGAAA	840
Qy	841	CGTTCAGTGCCAGTGACAAACGAAAGCATCAAGCATCCGTTTTCGAAAATGTTTACTGC	900
Db	841	CGTTCAGTGCCAGTGACAAACGAAAGCATCAAGCATCCGTTTTCGAAAATGTTTACTGC	900
Qy	901	CGGAGCGAAACAACTTCGAAAGCTACAGAAATGGGTAAATGGCAGAGCTGATGAAAAATCC	960
Db	901	CGGAGCGAAACAACTTCGAAAGCTACAGAAATGGGTAAATGGCAGAGCTGATGAAAAATCC	960
Qy	961	RACTGAACCTAGAAAGCACAAGAGAGTTAGACAAGTATTGTTGTAATGGGAAAGT	1020
Db	961	RACTGAACCTAGAAAGCACAAGAGAGTTAGACAAGTATTGTTGTAATGGGAAAGT	1020
Qy	1021	TGATGAATCAAGATTTTCATGATTTTGAAATTTCTTCAAGTTAGTGGTTAAGAACTCTAAG	1080
Db	1021	TGATGAATCAAGATTTTCATGATTTTGAAATTTCTTCAAGTTAGTGGTTAAGAACTCTAAG	1080
Qy	1081	ATTATCATCTCCGGTGTCTTGAATTCGAGGAGGTGTAGAGAAAACAACGAAATTTGATGG	1140
Db	1081	ATTATCATCTCCGGTGTCTTGAATTCGAGGAGGTGTAGAGAAAACAACGAAATTTGATGG	1140
Qy	1141	ATATGAATTTATCCGACACATCGAATGTTCTGATGCTTCGGCGATAGGAGAGATCC	1200
Db	1141	ATATGAATTTATCCGACACATCGAATGTTCTGATGCTTCGGCGATAGGAGAGATCC	1200
Qy	1201	TAAATCTTGGTCCGAACTCGGAAAGTTTAAACCCAGAAAGGTTTAAAGATTTGTGCAATTTGA	1260
Db	1201	TAAATCTTGGTCCGAACTCGGAAAGTTTAAACCCAGAAAGGTTTAAAGATTTGTGCAATTTGA	1260
Qy	1261	TTATAAAGGACGACATTTTGAACCTGGTACCATTTGGTGGTGCAGGAAAAAGAAATATGCTCTGG	1320
Db	1261	TTATAAAGGACGACATTTTGAACCTGGTACCATTTGGTGGTGCAGGAAAAAGAAATATGCTCTGG	1320
Qy	1321	CATTACTTCAGCTATTTACCAATTTGGAGTATGTCATTATAAATCTATTATATCATTTTAA	1380
Db	1321	CATTACTTCAGCTATTTACCAATTTGGAGTATGTCATTATAAATCTATTATATCATTTTAA	1380
Qy	1381	TTGGAACTGGCCGATGGAATTTACACCTCAAACTTGATATGACTGAAGCTATTGGCGG	1440
Db	1381	TTGGAACTGGCCGATGGAATTTACACCTCAAACTTGATATGACTGAAGCTATTGGCGG	1440
Qy	1441	TGCTCTCAGGAAAAAATAGATCTTAAGTTGATTCCTATTCATATCAAGTTAGCTTAGG	1500
Db	1441	TGCTCTCAGGAAAAAATAGATCTTAAGTTGATTCCTATTCATATCAAGTTAGCTTAGG	1500
Qy	1501	CTCAAAATATTTCTTGATTACATAGAGGTTTGAATATATATAATAAACTTTAAATAACG	1560
Db	1501	CTCAAAATATTTCTTGATTACATAGAGGTTTGAATATATATAATAAACTTTAAATAACG	1560
Qy	1561	ATGTTCTAATAATGTTTGGGTAGTTAATAAGTTTTCACCCGATCATATAAGTAGGCT	1620
Db	1561	ATGTTCTAATAATGTTTGGGTAGTTAATAAGTTTTCACCCGATCATATAAGTAGGCT	1620
Qy	1621	TCCTTGTAGGATGGTTAGATTATAAGTGTGGTGGTGGATTTTATAGATGGGTTAAAT	1680
Db	1621	TCCTTGTAGGATGGTTAGATTATAAGTGTGGTGGTGGATTTTATAGATGGGTTAAAT	1680
Qy	1681	GATTGGATGGAATAATAATAAATGAAATGTTTCTTTTCCAAATCCGAAAAA	1733
Db	1681	GATTGGATGGAATAATAATAAATGAAATGTTTCTTTTCCAAATCCGAAAAA	1733
RESULT 2			
AF406732			
LOCUS	AF406732	1758 bp	linear
DEFINITION	Euphorbia lagascae cytochrome P450 (CYP726A1)		PLN 13-FEB-2002
ACCESSION	AF406732.1		
VERSION	AF406732.1		
KEYWORDS	GI:18157658		
SOURCE	Euphorbia lagascae		
ORGANISM	Euphorbia lagascae		
REFERENCE			
AUTHORS	Cahoon, E.B., Ripp, K.G., Hall, S.E. and McGonigle, B.		
TITLE	Transgenic production of epoxy fatty acids by expression of a		
JOURNAL	cytochrome P450 enzyme from Euphorbia lagascae seed		
PUBMED	Plant Physiol. 128 (2), 615-624 (2002)		
REFERENCE	1 (bases 1 to 1758)		
AUTHORS	Cahoon, E.B.		
TITLE	Direct Submission		
JOURNAL	Submitted (06-AUG-2001) DuPont Crop Genetics, E.I. DuPont Company,		
FEATURES	Location/Qualifiers		
source	1..1758		
gene	/organism="Euphorbia lagascae"		
CDS	/mol_type="mRNA"		
	/db_xref="taxon:54672"		
	/gene="CYP726A1"		
	1..1758		
	/note="delta12-fatty acid epoxigenase"		
	21..1523		
	/product="cytochrome P450"		
	/codon_start=1		
	/protein_id="AAL62063.1"		
	/db_xref="GI:18157658"		
	/translation="MEQKNLSFPSPSILISFLVLILVVMRLWKKNPPPPKPFPIIG		
	NLPHLLLTSDLGHERFRALQITGVPMSLQIGVSAVVISSAAAEKVNKMTQADAPAQ		

RPVLDQAIVFYNRKDVLFASVCDHWRQMKXWILFLSACKVOSSRLIREEWEDJAL
FLTSRASKSPNITKILGYLISIMRTSVGNCKQRKLLSPADNAVEAATSFCTQDAN
FPFWLHLYHIAESKPRUHQEIDILIEENKHKANPEADNMLVILNQKGN
VPVPTENISKASVQMFQTAGESTTSKATWMAEMKPNITELSAQEVRFVGEFNG
KVDESRHDIKFLFKVETLSTAGETTSKATWMAEMKPNITELSAQEVRFVGEFNG
GRDPTNWSPEKGNPRFKDCAIDYKGTFTFELVPGAGKRCPGTISATINLEVIN
LHLYFNWELADGTPOTDWTGAJGALKRKIDKLPIYPTOVSLSGISV

ORIGIN

```
Query Match      99.8%; Score 1729.8; DB 4; Length 1758;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1731; Conservative 0; Mismatches 2; Indels 0;
```

Qy	1	GCATAAAGGAAAATGAGCAGAAAAATCTCTCTCTTTCCGAGCAATTTTAAATAGTTTTCT	60
Db	8	GCAGAAAGGAAAATGGAGCAGAAAAATCTCTCTTTCCGAGCAATTTTAAATAGTTTTCT	67
Qy	61	GCCTGTTTTTAATCTTAGTAGTAGTCATGAGTGTGTGGAAGAAACAGAAATCCACTCCAGG	120
Db	68	GCCTGTTTTTAATCTTAGTAGTAGTCATGAGTGTGTGGAAGAAACAGAAATCCACTCCAGG	127
Qy	121	GCCATGGAAGTTTCCTATCATAGGTAATCTTCCCTCATTTATTACTCACTTCTGATCTAGG	180
Db	128	GCCATGGAAGTTTCCTATCATAGGTAATCTTCCCTCATTTATTACTCACTTCTGATCTAGG	187
Qy	181	CCATGAAAGCTTTAGAGCCTTGGCTCAAAATTTATGGACCTGTTATGAGTCTTCAAAATGG	240
Db	188	CCATGAAAGCTTTAGAGCCTTGGCTCAAAATTTATGGACCTGTTATGAGTCTTCAAAATGG	247
Qy	241	CCAAGTTTCAGCTGTTGTCTCAATTTCTTCCAGCTGAAGCAGCCAAAGAGTTATGAAACTCA	300
Db	248	CCAAGTTTCAGCTGTTGTCTCAATTTCTTCCAGCTGAAGCAGCCAAAGAGTTATGAAACTCA	307
Qy	301	GGCTGATGCCCTTGGCCCAAAGCCCTATCGCTTGGACGCACAGATTCGTGTTTTATTAATCG	360
Db	308	GGCTGATGCCCTTGGCCCAAAGCCCTATCGCTTGGACGCACAGATTCGTGTTTTATTAATCG	367
Qy	361	GAAGAATGTCCTGTTGTCCTTCATATGGAGATCACTGGAGGCAGATGAAGAAAAATTTGGAT	420
Db	368	GAAGAATGTCCTGTTGTCCTTCATATGGAGATCACTGGAGGCAGATGAAGAAAAATTTGGAT	427
Qy	421	ACTTGAAATTTTCAGTGCCTGCAAAAAGTTCAATCTCTCAGGTTAATCCGAGAGGAAGAAAT	480
Db	428	ACTTGAAATTTTCAGTGCCTGCAAAAAGTTCAATCTCTCAGGTTAATCCGAGAGGAAGAAAT	487
Qy	481	GGAGGATGCCATCACATTTCTCCGTTTCGAAGCCGGATCTCCGGTCAATATTAACAAGAT	540
Db	488	GGAGGATGCCATCACATTTCTCCGTTTCGAAGCCGGATCTCCGGTCAATATTAACAAGAT	547
Qy	541	CATTTATGGCATTTAATATTTTCGATCATGATAAGAAACATCCGTTGGTTAATTTGAAGCAAAA	600
Db	548	CATTTATGGCATTTAATATTTTCGATCATGATAAGAAACATCCGTTGGTTAATTTGAAGCAAAA	607
Qy	601	AGAAAGATTGCTGAGTGTTCGGATCGAGTCAATGAGGCAGCAGGTTTTGCGCACCGC	660
Db	608	AGAAAGATTGCTGAGTGTTCGGATCGAGTCAATGAGGCAGCAGGTTTTGCGCACCGC	667
Qy	661	AGACGCTTTTCCGACGTGGAAATTACTTCACTATATCATTTGGAGCTGAGTCAAAACCCAG	720
Db	668	AGACGCTTTTCCGACGTGGAAATTACTTCACTATATCATTTGGAGCTGAGTCAAAACCCAG	727
Qy	721	GCCTTTGCATCAGAGATTCAAGATATACATTGGAAGAGATTCCTTAATGAACACAAAGCCAA	780
Db	728	GCCTTTGCATCAGAGATTCAAGATATACATTGGAAGAGATTCCTTAATGAACACAAAGCCAA	787
Qy	781	TAAAGCCTTTTGAAGCCGATTAACCTTAATGGATGTTCTATTGAATCTTCAAAAAAATGGAAA	840
Db	788	TAAAGCCTTTTGAAGCCGATTAACCTTAATGGATGTTCTATTGAATCTTCAAAAAAATGGAAA	847
Qy	841	CGTTCCAGTCCGAGTGCAAAACGAAAGCATCAAGACATCCGTTTGGCAAAATGTTTACTGC	900
Db	848	CGTTCCAGTCCGAGTGCAAAACGAAAGCATCAAGACATCCGTTTGGCAAAATGTTTACTGC	907

[illegible]

RESULT 3
CS223906

LOCUS	CS223906	1538 bp	DNA	linear	PAT 15-DEC-2005
DEFINITION	Sequence 162 from Patent WO2005111217.				
ACCESSION	CS223906				
VERSION	CS223906.1	GI:83685434			
KEYWORDS					
SOURCE	Nicotiana tabacum (common tobacco)				
ORGANISM	Nicotiana tabacum				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; famlids; Solanaceae; Solanaceae; Nicotiana.				
REFERENCE	1				
AUTHORS	Xu, D.				
TITLE	Nicotiana nucleic acid molecules and uses thereof				
JOURNAL	Patent: WO 2005111217-A 162 24-NOV-2005;				

```

FEATURES
  source
    U.S. Smokeless Tobacco Company (US)
    Location/Qualifiers
      1..1538
      /organism="Nicotiana tabacum"
      /mol_type="unassigned DNA"
      /db_xref="taxon:4097"
ORIGIN
  Query Match      23.7%; Score 410; DB 2; Length 1538;
  Best Local Similarity 58.1%; Pred. No. 5e-116;
  Matches 816; Conservative 0; Mismatches 555; Indels 33; Gaps 4;
QY 113 CCTCAGGGCCATGGAAGTTTCCTATCATAGTAACTCTCTCTCATCTTATTTACTCATTCT 172
DB 129 CCACCAGGTCCATGGAAACTACCAATACATAGGAAGTATGCTTCATATGGT-----TGGT 182
QY 173 GATCTAGGCCATGAACGTTTATAGAGCCTTGGCTCAAAATTTATGGACCTGTTATGAGTCTT 232
DB 183 GGACTACACACCATGTCCTTAGAGATTAGCCAAATAATATGGACCGCTTATGCACCTT 242
QY 233 CAAATGGCCCAAGTTTCAGCTGTGTCTATCTTCTCAGCTGAAGCAGCCAAAGAGGTTATG 292
DB 243 CAAATAGGTGAAGTTTTCGAGTTTGTGTTACTTCTCTGATATGSCAAAGAAAGTACTA 302
QY 293 AAACTCAGGCTGATGCCTTCGCCCAACGCCCTATCGTCTTGGAGCGACAGATTGTGTTT 352
DB 303 AAACTCATGACATCGCTTTCGCGTCTAGGCGCTAGCCCTTTTGGCCCCGGAGATTGCTGT 362
QY 353 TATAATCGGAAGATGTCTGTTTCTCATATGAGATCACTGGAGCGACAGTGAAGAAA 412
DB 363 TACAATAGCTGATCTTCGTTTTGCCCCCTATGCGAATTATGGAGACAAATGCGTAAA 422
QY 413 ATTTGGATACITGAATTTCTGAGTCCCAAAAAGTTCAATCTCCAGGTTAATCCGAGAG 472
DB 423 ATATGTGCTTGGAAAGTCTCAGTGCCCAAGAAATGTTGCGACATATAGCTCTATTAGGCGC 482
QY 473 GAAGAAATGGAGGATGCCATCACATTCCTCGTTGGAAGCCGATCTCCGGTCAATATT 532
DB 483 GATGAAGTTCTTCGTTCTCTTAATTTTATCCGGTCACTCTTCGTGGAGCCCTGTTAATATT 542
QY 533 ACAAGATCATTTATGGCATTAATAATTCGATCATGATAAGAAACATCCGTTGGTAA---T 589
DB 543 ACGGAAGATCTTTTGTTCACAGCTCCATGACATGTAGATCAGCGTTTGGGCAAGTA 602
QY 590 TGTAAAGCAAAAAGAAAGATTGCTGAGTGTGCGGATGCAATGAGCGACGACGAGT 649
DB 603 TTCAAAGGAGCAAGCAAAATTTATACAACTAATTAAGAAGTTATATCTCTTAGCAGGAGG 662
QY 650 TTTGSCACCGCAGACGCTTTTCGAGCTGGAATTTACTTCACTATATCATTTGGAGCTGAG 709
DB 663 TTTGATGTGGCTGACATATTCCCTTCATACAAAGTCTCTTCATGTGCTCAGTGAATGAAG 722
QY 710 TCAAAACCCAGGCGTTTGATCAGGAGATTGACGATATATCTTGAAGAGATTCTTAAATGAA 769
DB 723 GGTAGATTATGAATGACACACCATTAAGGTAGATGCTATTGTTGAGNATGTCATCAACGAG 782
QY 770 CAC-----AAAGCCAAATAAGCCCTTTTGAAGCGGATAACTTAATG 808
DB 783 CACAAGAAAAATCTTTGCAATTTGGAAAACTAATAGGAGCGTTAGGAGGTGAAGATTTAATT 842
QY 809 GATGTTCTATTGAATCTTCAAAAATGGAACCGTTCCAGTCCGAGTGAACCAAGCAAGC 868
DB 843 GATGTTCTTCTAAAATCTTATGAATGATGAGGAGCCCTTCAATTTCCCTATCAACCAAGCAAC 902
QY 869 ATCAAAGCATCGGTTTTCGAAATGTTTACTGCGGAGCGGAACAACTTCGAAAGCTACA 928
DB 903 ATCAAGCTATATCTTTGACATGTTTGTGCTGGAACAGAGACTTCATCGTCAACAATT 962
QY 929 GAATGGGTAAATGGCAGAGCTGATGAAAAATCCAACTGAACCTGAAGAAAGCAACAAGAA 988
DB 963 GTGTGGGCTATGTTGGAAATGGTGAATAATCCAACTGTATTTGCGAAAGCTCAAGCAGAA 1022
QY 989 GTTAGACAAGTATTTGGTGAATGGGAAAGTTGATGAATCAAGATTTTCATGATTGAAA 1048
1023 GTAAGAGATGCAATTTAGAGAAAAAGAAACCTTTTGATGAAATGATGTGGAGGAGCTAAAC 1082
1049 TTCTTCAAGTTAGTGGTTTAAAGAAACTCTAAGATTACATCTCCGGTT---GTCCTTGATT 1105
1083 TATCTAAAGTTAGTCAITTAAGAAACTCTAAGACTTTCATCCACCGGTTCCACTTTTGCTC 1142
1106 CCGAGGAGTGTAGAGAAAACAACAGCAATTGATGGATATGAANTTCATCCGAACTCGA 1165
1143 CCAAGAGAATGTAGGGAAGAGACAAATATAAACCGCTACACTATTCCTGTAAAGACCAAA 1202
1166 ATTCTGTGAATGCTTGGCGGATAGGAAGAGATCCTTAATCTTGTGTCGGAACCTCGGAAG 1225
1203 GTCAATGTTAATGTTTGGCAATTCGGAAGAGATCCAAATATTTGGATGATGCAGAACT 1262
1226 TTTAACCCAGAAAGGTTTAAAGATTGTGCAATTTGATTATATAAGGACGACATTTGAACTG 1285
1263 TTTAAGCCAGAGAGATTGAGCAGTGTCTAAGGATTTTGTGTAATAATTTTGAATAT 1322
1286 GTACATTTGGTCAGGAAAGAAATATGTCCTCGCATTTCTTCAGCTATTACCAATTTG 1345
1323 CTTCCATTTGGTGTGGAAGGAGATTGTCCAGGATTTCTGTTGGTTTACCTAATGCT 1382
1346 GAGTATGTCATTATAAATCTATTATATCATTTTAAATTTGGGAACCTGGCCGATGGAATTACA 1405
1383 TATTTGCCATTGGCTCAATTACTTTATCATTGTTGGAACCTCCCACTGGAATCAAA 1442
1406 CCTCAACACATTTGATATGACTGAAGCTATTGGCGGTGCTCTCAGGAAAAAATATAGATCTT 1465
1443 CCAAGCGACTTTGAGCTTGACTGAGTTGGTTGGAGTAAGTCCGCTAGAAAAAGTGACCTT 1502
1466 AAGTTGATTCCTATTCCATATCAA 1489
1503 TACTTGGTTCGACTCCTTTATCAA 1526
RESULT 4
CS223962 1673 bp DNA linear PAT 15-DBC-2005
LOCUS Sequence 218 from Patent WO2005111217.
DEFINITION CS223962
ACCESSION CS223962
VERSION CS223962.1 GI:83685860
KEYWORDS
SOURCE Nicotiana tabacum (common tobacco)
ORGANISM Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Nicotiana.
REFERENCE
1
  Xu,D.
  Nicotiana nucleic acid molecules and uses thereof
  Patent: WO 2005111217-A 218 24-NOV-2005;
  U.S. Smokeless Tobacco Company (US)
  Location/Qualifiers
    1..1673
    /organism="Nicotiana tabacum"
    /mol_type="unassigned DNA"
    /db_xref="taxon:4097"
FEATURES
  source
ORIGIN
  Query Match      23.1%; Score 400.4; DB 2; Length 1673;
  Best Local Similarity 57.7%; Pred. No. 5e-113;
  Matches 810; Conservative 0; Mismatches 561; Indels 33; Gaps 4;
QY 113 CCTCAGGGCCATGGAAGTTTCCTATCATAGTAACTCTCTCTCATCTTATTTACTCATTCT 172
DB 135 CCACCAGGTCCATGGAAACTACCAATACATAGGAAGTATGCTTCATATGGT-----TGGT 188
QY 173 GATCTAGGCCATGAACGTTTATAGAGCCTTGGCTCAAAATTTATGGACCTGCTTATGAGCTTT 232
DB 189 GGACTACCAACCATGCTCTTAGAGATTAGCCAAATAATATGGACCACTTATGCACCTT 248
233 CAAATTTGGCAAGTTTTCAGCTGTTGTCATTTCTTCAGCTGAAGCAAGCAAGGTTATG 292
```


Db 249 CAAATAGGTGAAAGTTTCTCGGTTGGTGTACTTCTCTGATACGCAAAAGAAATATTA 308
Qy 293 AAAAATCAGGCTGATGCGCTTCGCCAACGCCCTATCGTCTTGCGACGACAGATGTTGTTT 352
Db 309 AAAAATCAGATCGCTTTGGGCTTAGCCCTTAGCCCTTTGGCCCGGAGATGTCGT 368
Qy 353 TATAATCGGAAAGATGTCTTGTGTTTCATATGGAGATCACTGGAGGAGATGAAGAA 412
Db 369 TACAATAGTCTGATCTAGCTTTTGGCCCTATGGCGACTATTGGAGACAAATGCGTAA 428
Qy 413 ATTTGGATCTTGAATTTCTGAGTGCCAAAAGTTCAATCTCTCAGGTTTAATCCGAGAG 472
Db 429 ATATGTGCTTGGAAAGTGTCTAGTGCCAAAGATGTTGCGACATTTAGCTCTATTAGCGG 488
Qy 473 GAAGAAATGGAGGATGCCATCACATCTCTCGGTTGGAAAGCGGATCTCCGGTCAATATT 532
Db 489 ATGAAGTCTTCGCTCAATTAATTTATCCGGTCACTCTTGGTGAACCTAATATGTT 548
Qy 533 ACAAGATCAATTTATGGCATTATAATTTTCGATCATGATAAGAACATCCGTTGGTAA--T 589
Db 549 ACGGAAAGGATCTTTTGTTCACAGCTCCATGACATGTAGATCAGCGTTTGGCGAAGTG 608
Qy 590 TGTAAAGAAAGAAAGATGCTGAGTGTTCGCGATGCGATGCAATGAGGACGCGAGT 649
Db 609 TTCAAAGAGCAAGACAAATTTATACAACTAATTAAGAAAGTGATCTCTTAGCAGGAGG 668
Qy 650 TTTGSCACCGCAGAGCTTTTCGCGAGTGGAATTTACTTCACTATATCATTTGAGCTGAG 709
Db 669 TTTGATGTGGTGACATATTCCTCTCACTGAAGTTTCTTCATGCTCAGTGGAAATGAAG 728
Qy 710 TCAAAACCCAGCGGTTTGCATCAGGAGATTTGACGATATATCTTGAAGAGATTTCTTAATGAA 769
Db 729 GGTAAAGATTATGAATGCACACCATTAAGGTAGTGCCATTTGTTGAGAAATGTCATCATGAG 788
Qy 770 CAC-----AAAGCAAATAGCCCTTTTGAAGCGGATCACTTAATG 808
Db 789 CACAAGAAAAATCTTGGCAATTTGGGAAAACTAATGGAGCGTTAGGAGGTGAAGATTTAATT 848
Qy 809 GATGTTCTATTGAATCTTCAAAAAATGGAACGTTCCAGTCCAGTGACAAACCAAGC 868
Db 849 GATGTTCTTGAAGACTTATGAATGATGAGGCGCTTCAATTTCTTATCACCACGACAC 908
Qy 869 ATCAAAGCATCGTTTTCGCAATGTTTACTGCGGAGCGAAACAACTTCGAAAGCTACA 928
Db 909 ATCAAAGCTATAATTTTGACATGTTTGTGCGCGGACAGAGACTTCATCGTCAACAATT 968
Qy 929 GAATGGGTAAATGGCAGAGCTGATGAATAATCCAACTGAACCTAAGAAAAAGCAAGAGAA 988
Db 969 GTGTGGCTATGTTAGAAATGGTGAATAATCCAGCGGTATTCGCGAAAGCTCAAGCAGAA 1028
Qy 989 GTTACAGATATTGTTGTAATGGGAAAGTTGATGATCAAGATTTCAATGATTTGAA 1048
Db 1029 GTAAGAGAGCAATTTAGAGGAAAGAAACTTTTCGATGAAATGATGTGGAGGAGCTAAC 1088
Qy 1049 TTCTTCAAGTTAGTGGTTAAAGAAACTCTAAGATTACATCTCCCGTT---GTCTTGATT 1105
Db 1089 TACCTAAGTTAGTAATAAAGAAACTCTAAGACTTCACTCCACCGGTTCCACTTTTGCTC 1148
Qy 1106 CCGAGGAGGTAGAGAAACAACAAGATTGATGATGAAATTCATCCGAACTCGA 1165
Db 1149 CCAAGAGAATGTAGGGAAGAGACAAATATAAAGCGCTACACTATTTCTGTAAAGACCAA 1208
Qy 1166 ATTTGTTGATCTTGGCGATAGAGAGATCTTAATCTTGTGTCGAACTGGAAG 1225
Db 1209 GTCTGGTTAATGTTTGGGCTTTGGGAAGAGATCCAAAATATTTGGAATGACGCAAACT 1268
Qy 1226 TTTAAACCCAGAAAGGTTTAAAGATTGTGCAATTTGATTAATAAGGAGCAGACATTTCGAACTG 1285
Db 1269 TTTATGCCAGAGATTTGAGCAGTGTCTTAGGATTTTGTGTTAATAATTTGATAT 1328
Qy 1286 GTACCATTTGGTGCAGGAAAAAGAAATATGTCTCGCATTTACTTTCAGCTATTACCAATTTG 1345

Db 1329 CTTCCATTTGGTGGGGAAGGAGGATTTGCTCGGATTTCTGTTGGCTAGCTAATGCT 1388
Qy 1346 GAGTATGTCATTATAAATCTATTATATCAATTTAATTGGAACTGCCGATGGAATTACA 1405
Db 1389 TATTTGGCAATGGCTCAATTTACTATATCACTTCGATTTGAAACTCCCTGCTGGAATCGAA 1448
Qy 1406 CTTCAACACACTTGATATGACTGAAGCTATTGGCGGTGCTCTCAGGAAAAAATAGATCTT 1465
Db 1449 CCAAGCGACTTGACTTGACTGAGTTGGTGAGTAATTCGCCCTAGAAAAAGTGACCTT 1508
Qy 1466 AAGTTGATTTCTATTCCATATCAA 1489
Db 1509 TACTTTGGTGGCACTCTTATCAA 1532

RESULT 5
LOCUS CS223964 1673 bp DNA linear PAT 15-DEC-2005
DEFINITION Sequence 220 from Patent WO2005111217.
ACCESSION CS223964
VERSION CS223964.1 GI:83685861
KEYWORDS
SOURCE Nicotiana tabacum (common tobacco)
ORGANISM Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Nicotiana.

REFERENCE 1
AUTHORS Xu,D.
TITLE Nicotiana nucleic acid molecules and uses thereof
JOURNAL Patent: WO 2005111217-A 220 24-NOV-2005;
U.S. Smokeless Tobacco Company (US)
FEATURES
Location/Qualifiers
1..1673
/organism="Nicotiana tabacum"
/mol_type="unassigned DNA"
/db_xref="taxon:4097"

ORIGIN
Query Match 23.1%; Score 400.4; DB 2; Length 1673;
Best Local Similarity 57.7%; Pred. No. 56-113;
Matches 810; Conservative 0; Mismatches 561; Indels 33; Gaps 4;

Qy 113. CTTCCAGGCGCATGGAAGTTTCTCTATCATAGTAATCTTCTCTCATTTTACTCACTTCT 172
Db 135 CCACAGGCTCCATGGAACACTACCATCTAGGAAGTATGCTTCATATGCT-----TGT 188
Qy 173 GATCTAGGCCATGAAGCTTTTAGAGCTTTGGCTCAAAATTTATGGACCTGTTTATGAGTCTT 232
Db 189 GGACTTACCACACCATGTCTTAGAGATTTAGCAAAAAAATATGGACCACTTATGCACCTT 248
Qy 233 CAATTTGGCCAAAGTTTTCAGCTGTTGTCTATTTCTTTCAGCTGAAGCAGCCAAAGGTTATG 292
Db 249 CAATTAGGTGAGTTTCTGCGGTTGTTGTTTCTCTCTGATACGCAAAAGAAAGTATTA 308
Qy 293 AAAACTCAGGCTGATGCGCTTTCGCCCAACGCGCTATCTGCTTGGACGCAAGATTTGTTT 352
Db 309 AAAACTCATGACATCGCTTTTGGCTTAGGCTAGCCCTTTTGGCCCGGAGATTTGCTGT 368
Qy 353 TATAATCGGAAAGATGTCTTGTGTTGCTTTCATATGAGATCACTGAGGCGAGATGAAGAAA 412
Db 369 TACAATAGTCTGATCTAGCTTTCGCGCTTATGGCGACTATTGGAGACAAATGCGTAAA 428
Qy 413 ATTTGGATCTTGAATTTCTGAGTGCCAAAAGTTTCAATCTCTCAGGTTAATCCGAGAG 472
Db 429 ATATGTGTTTGGAAAGTGTCTCAGTSCCAAGAAATGTTTCGACATTTAGCTCTATTAGCGG 488
Qy 473 GAAAGAAATGGAGGATGCCATCACATTTCTCGGTTTCAAAAGCGGATCTCCCGTCAATATT 532
Db 489 AATGAAGTCTTCTGCTCATTAATTTATCCGGTCACTCTTCTGGTGAACCTATTATGTT 548
Qy 533 ACAAAGATCATTTTATGGCATTATAATTTTCGATCATGATAAGAAACATCCGTTGGTAA---T 589

Db 549 ACGGAAAGGATCTTTTGTTCACAAGCTCCATGACATGTAGATCAGCGTTTGGGCAAGTG 608
Qy 590 TGTAAAGCAAAAGAAAGATTGCTGAGTCTTGCCGATGCAATGAGCGACGACGAGT 649
Db 609 TTCAAGAGCAAGACAATTTATACAACTAATTAAGAAAGTGATCTTTAGCAGGAGG 668
Qy 650 TTTGGCCCGCAGACGCTTTTCCGACGCTGGAAATTTACTTCACTATATCAATTTGGAGCTGAG 709
Db 669 TTTGATGGCTGACATATCCCTTCACTGAAGTTTCTTCACTGTCTCAGTGGAAATGAAG 728
Qy 710 TCAAAACCCAGCGCTTTGCATCAGGAGATTGACGATATACCTTTGAAGAGATTCTTAATGAA 769
Db 729 GGTAAAGATTATGAATGCAACCATAAAGGTAGATGCCATTGTGAGATGTCAATCAATGAG 788
Qy 770 CAC-----AAAGCCAAATAAGCCCTTTTGAAGCCGATAAATCTTAATG 808
Db 789 CACAAGAAAAATCTTGCAATTTGGGAAAACTAATGGAGCGTTAGGAGGTGAAGATTTAAT 848
Qy 809 GATGTTCTATTGAATCTTTCAAAAAAATGGAACCGTTCCAGTGCCAGTGACAAAACGAAGC 868
Db 849 GATGTTCTTAAAGACTTATGAATGATGGAGCGCTTCAATTTCCCTATCACCACGACACAC 908
Qy 869 ATCAAGCATCCGTTTTCGCAATGTTTACTGCCGGGCGGAAACAACTTCCGAAAGCTACA 928
Db 909 ATCAAGCCATAATTTTGGCATGTTTCTGCGGGGACAGAGCTTCATCGTCAACAAT 968
Qy 929 GAATGGGTAATGGCAGAGCTGATGAAAAATCCAACTGAACCTAAAGAAAAAGCACAAGAA 988
Db 969 GTGTGGGCTATGGTAGAATGGTGAATAATCCAGCCGTTATTCGGCAAGCTCAAGCAGAA 1028
Qy 989 GTTAGACAAGATTTTGGTGAAATGGGAAAAAGTTGTAGTAATCAAGATTTTCATGATTTGAAA 1048
Db 1029 GTAAGAGAAGCATTTAGAGGAAAAAGAACTTTTCGATGAAAAATGATGTGGAGGAGCTAAAC 1088
Qy 1049 TTTCTCAAGTTAGTGGTTTAAAGAACTCTAGATTTACATCTCCCGTTT- - -GTCTTGATT 1105
Db 1089 TACCTAAAGTTAGTAATAAAGAAACTCTAGAGCTTTCAACCCGTTCCACTTTTGTCT 1148
Qy 1106 CCGAGGAGTGTAGAGAAACAAACGAATTTGATGGATATGAAATTCATCCGAAACACTCGA 1165
Db 1149 CCAAGAGATGTAGGGAAGAGACAATATAACGGCTACACTATTTCTGTAAAGACCAA 1208
Qy 1166 ATTGTTGTAATGCTTTGGCGGATAGAGAGATCTTAATATCTTGTCGGAACCTGGAAG 1225
Db 1209 GTCATGTTAATGTTTGGGCTTTGGGAAAGAGATCCAAAAATATGGAATGACGACAGAACT 1268
Qy 1226 TTTAACCCAGAAAGCTTTAAAGATTGTGCAATTTGATTATAAGGACGACACATTTGAACTG 1285
Db 1269 TTTATGCCAGAGATTTGACAGTGTCTAAGGATTTTGTGGTAATTAATTTTGAATAT 1328
Qy 1286 GTACCATTTTGGTCGAGAAAAAGAAATATGTCCTGGCATTTACTTTCAGCTATTACCAATTTG 1345
Db 1329 CTTCCATTTTGGTGGCGGAAGGAGATTGTCCTGGGATTTGCTTGGCTTAGCTAATGCT 1388
Qy 1346 GAGTATGCAATTAATAATCTATATATCAATTTTAAATTTGGAACTGGCGGATGGAATTACA 1405
Db 1389 TATTTGCCATTGGCTCAATTAATATATATCACTTCGATTTGGAATCTCCCTGCTGGAATCGAA 1448
Qy 1406 CCTCAACACTTGATATGACTGAAGCTATTGGCGGTGCTCTCAGGAAAAAATAATAGATCTT 1465
Db 1449 CCAAGGCACTTGGACTGAGTGGTTGGTAGTAACCTGCGCTAGAAAAAAGTGACCTT 1508
Qy 1466 AAGTTGATTCCTATTCCATATCAA 1489
Db 1509 TACTTGGTTGCGACTCTTATCAA 1532

RESULT 6
LOCUS CS223984 1673 bp DNA linear PAT 15-DEC-2005
DEFINITION Sequence 240 from Patent WO200511217.
ACCESSION CS223984
VERSION CS223984.1 GI:83685871

KEYWORDS Nicotiana tabacum (common tobacco)
SOURCE Nicotiana tabacum
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Solanales; Solanaceae; Nicotiana.
REFERENCE 1
AUTHORS Xu,D.
TITLE Nicotiana nucleic acid molecules and uses thereof
JOURNAL Patent: WO 200511217-A 240 24-NOV-2005;
U.S. Smokeless Tobacco Company (US)
FEATURES
source Location/Qualifiers
1..1673
/organism="Nicotiana tabacum"
/mol_type="unassigned DNA"
/db_xref="taxon:4097"
ORIGIN
Query Match 23.1%; Score 400.4; DB 2; Length 1673;
Best Local Similarity 57.7%; Pred. No. 5e-113;
Matches 810; Conservative 0; Mismatches 561; Indels 33; Gaps 4;
Qy 113 CCTCAGGGCCATCGAAGTTTCCTATCATAGGTAATCTTCCTCATTTATTACTCATCTCT 172
Db 135 CCACAGGTCATGGAACCTACCAATACTAGGAAGTATGCTTCATATGCT-----TGCT 188
Qy 173 GATCAGGCCATGAACGTTTGTAGACCTTGGCTCAAATTTATGACCTGTTATGAGCTTT 232
Db 189 GGACTACCAACCATGCTCTTAGAGATTTAGCCAAAAAATATGGACCACTTATGCACCTT 248
Qy 233 CAAATTGGCCAAAGTTTCAGCTGTTGTCAATTTCTCAGCTGAAGCAGCCAAAGAGTTATG 292
Db 249 CAATTAGGTGAAGTTTCTGCGGTTGTGGTACTTCTCTGATACCGCAAAAGAGATTATA 308
Qy 293 AAAACTCAGGCTGATGCCCTTCGCCAACGCCCTATCGTCTTGGAGCAGAGATTGTGTTT 352
Db 309 AAACTCATGATCGCTTTTGGCTTAGGCTAGCCTTTGGCCCCGGAGATTGCTGT 368
Qy 353 TATAATCGAAAGATGCTCTGTTTGTTCATATGAGATCACTGGAGGAGATGAAGAAA 412
Db 369 TACAATAGTCTGATCTAGCCCTTTTGGCCCTATGGCGACTATTGGAGACAAATGCGTAAA 428
Qy 413 ATTTGGATCTTGAATTTCTGAGTGCCAAAAAGTTTCAATCTCCAGGTTAATCCGAGAG 472
Db 429 ATATGTGTCTTGGAGTGCTCAGTGCCAAAGATGTTTCGGAACATTTAGCTCTATTAGCGG 488
Qy 473 GAAAGATGGAGGATGCCATCACATTCCTCCGTTTCGAAAGCCGGATCTCCGTCATATT 532
Db 489 AATGAAGTTCTCGTCTCATTTAAATTTATCCGGTCACTCTCTGTTGAACCTATTAAATGT 548
Qy 533 ACAAGATCATTTATGGCATTTAATTTGATCATGATAAGAACATCCGTTGGTAA--T 589
Db 549 ACGGAAAGGATCTTTTGTTCACAAGCTCCATGACATGTAGATCAGCGTTTGGGCAAGTG 608
Qy 590 TGTAAAGCAAAAGAAAGATTGCTGAGTGTGCGGATGCAATGAGCAGCAGCAGT 649
Db 609 TTCAAGAGCAAGACAATTTATACAACTAATTAAGAAAGTGATCTCTTAGCAGGAGG 668
Qy 650 TTTGSCACCGCAGACGCTTTTCCGACGCTGGAATTTACTTCACTATATCAATTCAGGAGCTGAG 709
Db 669 TTTGATGTGGCTGACATATTCCTTTCACTGAAGTTTCTTTCATGTCTCAGTGGAAATGAAG 728
Qy 710 TCAAAACCCAGCGCTTTTCCGACGCTGGAATTTACTTCACTATATCAATTCAGGAGCTGAG 769
Db 729 GGTAAAGATTATGAATGCAACCATAAAGGTAGATGCCATTGTGAGATGTCAATCAATGAG 788
Qy 770 CAC-----AAAGCCAAATAAGCCCTTTTGAAGCCGATAAATCTTAATG 808
Db 789 CACAAGAAAAATCTTGCAATTTGGGAAAACTAATGGAGCGTTAGGAGGTGAAGATTTAAT 848
Qy 809 GATGTTCTATTGAATCTTTCAAAAAAATGGAACCGTTCCAGTGCCAGTGACAAAACGAAGC 868
Db 849 GATGTTCTTAAAGACTTATGAATGATGGAGCGCTTCAATTTCCCTATCACCACGACACAC 908

Db	986	CAAAGGCACAAGCTGAAGTGAGACAAGTCTTTGAAGGAAAAGAAAGGTTTTTCAACAATTTG	1045
Qy	1032	GATTTTCATGATTTGAAATTTCTTCAAGTTAGTGGTTTAAAGAAACTCTAAGATTATACATCTC	1091
Db	1046	ATCTTGATGAGTTGAAGTACTTTGAAGTTAGTAATCAAGAAACTCTAAGGATGCACCCCTC	1105
Qy	1092	CGGT- --TGCTTTGATTTCCGAGGAGTGTAGAGAAAACAACAGAAATTCATGATATGAAA	1148
Db	1106	CAATTCCTCTATTAGTCCCTAGAGAAATGTATGAAGGATACAAAGATTGACGGGTACAATA	1165
Qy	1149	TTCATCCGAAACACTCGAATTTGTTGTGAATGCTTTGGCGCATAGGAGAGATCCCTAATACTT	1208
Db	1166	TACCTTTCAAACTCGAGTCAATAGTTAATGCATGGGCAATTGGACGAGATCCCTGAAAGTT	1225
Qy	1209	GTCGGAACTGGAAGCTTTAAACCCAGAAAGTTTAAAGATTGTGCAATTTGATTATAAAG	1268
Db	1226	GGGATGACCTTGAAGCTTTTCCCGAGAGATTCGAGAAATAGTTCTGTGTGACTTTCTTGTG	1285
Qy	1269	GGACGACATTTGAACCTGGTACCAATTTGGTGCAGGAAAAGAAATATGTCTGGCATTACTT	1328
Db	1286	GAAGCCATCATCAATTTATTTCATTTGGTGGGGAAGAGATTTGTCTCGAATGCTTT	1345
Qy	1329	CAGCTATTACCAATTTGGAGTATGTCATATAAATCTATTATATCATTTTAAATTTGGAAC	1388
Db	1346	TTGGTTTAGCCAATTTGGCAACCATTAGCTCAATTTACTTTATCACTTGATCGGAAC	1405
Qy	1389	TGGCCGATGGAATTAACCTCAACACTTGATATGACTGAAGCTATTGGCGGTGCTCTCA	1448
Db	1406	TCCCTAATGGCAAAAGTCAGAAATTTGGACATGACGGAGTCACTTGGAAATTTCTGCA	1465
Qy	1449	GGAAAAAATAGATCTTAAAGTTGATTCCTATTCCATATCAAGTTAGCTTAGGCTCAAATA	1508
Db	1466	CAAGAAAGGATGATCTTGTGTTGATTCGACCCCTTATGATCCTTTGAATGTATTGAGACA	1525
Qy	1509	TTTCTTGATTACATGAGAGGTTGAAA	1535
Db	1526	GTTGTAGAAATAAAAAAGAGGGAGAA	1552
RESULT 8			
AR205821			
LOCUS	1781 bp DNA linear PAT 20-JUN-2002		
DEFINITION	Sequence 1 from patent US 6369212.		
ACCESSION	AR205821		
VERSION	AR205821.1 GI:21503501		
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 1781)		
AUTHORS	Oh,B.-J., Ko,M.Kyung, and Kim,Y.Soon.		
TITLE	Cytochrome P450 gene highly expressed in the incompatible		
interaction			
JOURNAL	Patent: US 6369212-A 1 09-APR-2002;		
FEATURES	Location/Qualifiers		
source	1..1781		
	/organism="unknown"		
	/mol_type="unassigned DNA"		
ORIGIN			
Query Match 23.0%; Score 399.4; DB 2; Length 1781;			
Best Local Similarity 56.3%; Pred. No. 1e-112;			
Matches 826; Conservative 0; Mismatches 611; Indels 30; Gaps 3;			
Qy	99	AGAAACAGAAATCCACTCCAGGGCCATGGAAGTTTCTCATCATAGTAAATCTTCTCTCAT	158
Db	86	AAAGCTAAACTTACCTCTCTGGTCCATGGAATATACCTTTTATTGGAAGCCTACATCACT	145
Qy	159	TATTACTCACTTCTGATCTAGGCCATGAACGTTTTAGAGCCTTGGCTCAAAATTTATGGAC	218
Db	146	TGGCAGTGGCAGTCCACCTTCTCATCATGCGCTAAAGAAATCTAGCAAACTTTATGGGC	205
Qy	219	CTGTTATGAGTCTTCAAAATTTGGCAAGTTTTCAGCTGTTGTGTCATTTCTTCAGCTGAAGCAG	278

Db	206	CGCTCATGCACTTACGACTCGGGAATTTCTCTACGTCATCATTTTCTCCCGGGAATGG	265
Qy	279	CCAAAGAGGTTTATGAAACTCAGGCTGATGCTTTGCGCCAAACGCCCTATCGCTTTGGACG	338
Db	266	CGAAGGAAGTACTAAANAACACACACCTCGCTTTTGCACACGAGCGGAAACTTGTGGTGG	325
Qy	339	CACAGATGTGTTTATTAATCGGAAAGATGCTGTTTGTCTTTCATATGGAATCACTCTGA	398
Db	326	CTGACATCGTCCATTAATAGTAGTAGCGGATATAGCATTTTCTCCATATAGTGAATACTGA	385
Qy	399	GGCAGATGCAAAATTTGGATCTTGAAATTTCTGAGTGCCAAAAAAGTTCATCTCTCA	458
Db	386	GGCAGATTCGTAATAATTTGCACTCTCGAACTCTCTTAGTGCCCAAGATGTCATCTTTA	445
Qy	459	GGTTAAATCCGAGAGGAAATGGAGGATGCCATCACATTCCTCGTTCGAAAGCCGAT	518
Db	446	GCTCAATTCGCCAGGATGAGCTGTCGATGATGGTCTCATCTATAGCAACCATGCCAATT	505
Qy	519	CTCCGGTCAATATTAATAAGATCAATTTATGCAATTAATTTTGCATCATGATAAGAACT	578
Db	506	TTCCCGTCAACCTTACAGACAAATAATTTTGGTTTACAAAGTTTCGGTAACTTGTAGATCAG	565
Qy	579	CCGTTGGTAA--TTGTAAGCAAAAGAAAGATGCTGAGTGTTCGCCGATGCAGTCAATG	635
Db	566	CTCTGGGAAAAATATGTCGTGACCAAGACAACTGATTAATTTTCATGAGGAAATAATAT	625
Qy	636	AGGCAGCCACGAGTTTGGCACCCGACAGCGCTTTTCCGACGTGGAAATTAATCTTCACTATA	695
Db	626	CATTGACAGGTGGATTTAGTATTCGCTGATTTTCCCTACATGGAATAATGCTACATGATG	685
Qy	696	TCAITGGAGCTGAGTCAAAACCCAGGCGTTTGCATCAGGAGATTCACCATATACTTGAAG	755
Db	686	TTGGTGGTTCAAAAACCTAGACTGCTGAAAGGCTCATCGTAAATTCGATGAGATTTTGGAA	745
Qy	756	AGATTCTTAATGAACACAAAGCCAAATAGCCCT-----TTTG	791
Db	746	ATGTAGTAATGAGCACAAACAGAAATCGAGCGGATGGCCAAAGGGTAATGCGGAATTTG	805
Qy	792	AAGCGGATAACTTAATGGATGTTCTATTGAATTTTCAAAAAAATGGAAACGTTTCAGTGCC	851
Db	806	CGGTTGAAGATTTGATCGATGTTTGTGTAAGGTTTCGAGAAAGTGGAGAAAGTTCAAAAT	865
Qy	852	CAGTGACAAACGAAAGCATCAAGCATCCGTTTTCGCAATGTTTACTCCGCGAGCGAAA	911
Db	866	CCATCACGGATGACAAATATCAAAATCAATATTAGTGGACATGTTCTCCGCTGGATCTGAAA	925
Qy	912	CAACTTCGAAAGCTACAGAAATGGGTAATGGCAGAGCTGATGAAAAATCCAACTGAACATA	971
Db	926	CGTATCGCAACTATATAATTTGGGCATTTAGCTGAAATGATGAAGAAACCAAGTGTCTAG	985
Qy	972	GAAAGCAACAAGAAAGTTAGACAAGTATTGTTGGTGAATGGGAAATGGGAAAGTTGATGAATCA	1031
Db	986	CAAGGGCAACAGCTGAAGTGAGACAAGCTTGAAGGAAAGAAAGGTTTTCACAAATTTG	1045
Qy	1032	GATTTTCATGATTTGAAATTTCTCAAGTTAGTGGTTTAAAGAACTCTAAGATTACATCTCTC	1091
Db	1046	ATCTTGATGAGTTGAAGTACTTGAAGTTAGTAATCAAGAAACTCTAAGGATGCAACCTCT	1105
Qy	1092	CGGT---TGTCTTGATTCGAGGGAGTGTAGAGAAACCAACGAAATTTGATGATATGAAA	1148
Db	1106	CAATTCCTCTATTAGTCCCTAGAGAAATGTATGAAGGATACAAAGATTGACGGGTACAAATA	1165
Qy	1149	TTTATCCGAAACACTCGAATTTGTTGAAATGCTTGGGGCGATAGGAAGAGATCTCTAATCTT	1208
Db	1166	TACCTTTCAAACTCGAGTCAAGTAAATTAATGATGCGGCAATTTGGACGAGATCTCTGAAGTT	1225
Qy	1209	GGTCCGAACCTCGAAAGTTTAAACCCAGAAAGGTTTAAAGATTTGCAATTTGATTAATAAG	1268
Db	1226	GGGATGACCCCTGAAAGCTTTTCCCGAGAGAGATTCGAGAAATAGTTCTGTTGACTTCTTG	1285
Qy	1269	GGACGACATTTGAACTGGTACCATTTGGTGGCAGGAAAAAGAAATATGTCTCTGCAATTTT	1328

Db 1286 GAAGCCATCATCAATTTATTCCTTTGGTGGGGAAGAAGATTGTCCTGGAATGCTTT 1345
QY 1329 CAGCTATTACCAATTTGGAGTATGTCATTATTAATCTATTATATCAATTTTAATGGGAAC 1388
Db 1346 TTGGTTAGCCAATGTTGGACAACCATGCTCAATTTACTTTTATCACTTCGATCGGAAAC 1405
QY 1389 TGCCGATGGAATTACACCTCAACACCTTGATATGATGAGTGAAGCTATTGGCGGTGCTCTCA 1448
Db 1406 TCCCTAATGACAAAGTCACGAAATTTGGACATGACGAGTACACCTGGAAATTTTGCA 1465
QY 1449 GGAATAAATAGATCTTAAGTTGATTCCTATTCCATATCAATCAAGTAGCTTAGGCTCAATA 1508
Db 1466 CAAGAAGGATGATCTGTTGTTGATGGCCACCTTATGATCCTTGAATGATTGAGACA 1525
QY 1509 TTTCTTGTATTACATAGAGGGTTGAAA 1535
Db 1526 GTTGTAGAAATAAAGAAGGGAGAA 1552

RESULT 9
AF122821 1781 bp mRNA linear PLN 08-FEB-2000
LOCUS Capsicum annuum cytochrome P450 (PepCYP) mRNA, complete cds.
DEFINITION AF122821
ACCESSION AF122821
VERSION AF122821.1 GI:6739505
KEYWORDS
SOURCE Capsicum annuum
ORGANISM Capsicum annuum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiales; Solanales; Solanaceae; Capsicum.

REFERENCE 1 (bases 1 to 1781)
Oh, B.J., Ko, M.K., Kim, Y.S., Kim, K.S., Kostenyuk, I. and Kee, H.K.
A cytochrome P450 gene is differentially expressed in compatible
and incompatible interactions between pepper (Capsicum annuum) and
the anthracnose fungus, Colletotrichum gloeosporioides
Mol. Plant Microbe Interact. 12 (12), 1044-1052 (1999)
10624013

REFERENCE 2 (bases 1 to 1781)
Oh, B.J., Ko, M.K., Kostenyuk, I., Kim, Y.S., Kim, K.S. and Kee, H.K.
Direct Submission
Submitted (22-JAN-1999) Kumbo Life and Environmental Science
Laboratory, Kumbo Petrochemical Co., LTD., 1 Oryong-dong, Puk-gu.,
Kwangju 500-712, Korea

FEATURES
Location/Qualifiers
1..1781
/organism="Capsicum annuum"
/mol_type="mRNA"
/db_xref="taxon:4072"
gene 1..1781
/gene="PepCYP"
CDS 4..1512
/gene="PepCYP"
/codon_start=1
/product="cytochrome P450"
/protein_id="AAF27282.1"
/db_xref="GI:6739506"
/translation="MEIQTNLVAFLLFLLSIILLKKWTKLNLPFGPKLPRIGS
LHLAVAGLPFHGLNKLKXGPLMHLRGEIPVTIISPRMAKEVLKTHLAFATR
PKLVADI VHYDSTDIATFSGYEWIRKIKICILELSAKMVKFFSSIRQDLSMMVS
SIRTPMFPNVTDKI FWFSTSVTCRSALGKICRDQDKLII FMR EII SUTGFSIADF
FTWKMLHDVGSKTRLLKAHRKIDILEHVNHKONRAGKNGEFGGBDLIDLVL
LVRSGEVIQISITDNIKSIIVDMFSGSETSTTIWALAEMKKPSVLAKADEV
ROVLKEKGFQIDLDLKYLVIKETLRMHPPILPLVPRCKDKIDGYNIPKPT
RVIVNAWAIGRPDSESPFSESPFSESVDFLGSHHQIFPFAGRRICPGMLFGL
ANVGPLAQLLYHFDRKLPNGOSHENDMTESPGISATRKDDLVLIATPYDP"

ORIGIN

Query Match 23.0%; Score 399.4; DB 4; Length 1781;
Best Local Similarity 56.3%; Pred. No. 1e-112;
Matches 826; Conservative 0; Mismatches 611; Indels 30; Gaps 3;
QY 99 AGAAACAGAAATCCACCTCCAGGCGCATGGAAGTTTCTCATCATAGGTAATCTTCTCTCAT 158

Db 86 AAAAGCTAAACTTACCTCTCGTCCATGGAATATTACTTTTATGGAAGCTACATCACT 145
QY 159 TATTACTCACTTCTGATCTAGGCCATGAACGTTTTAGAGCCTTGGCTCAAAATTTATGGAC 218
Db 146 TGGCAGTGGCAGTCCACTTCTCATCATGCGCTAAAAAATCTAGCAAAATTTATGGGC 205
QY 219 CTGTTATGAGTCTTCAAAATGGCCAAAGTTTTCAGCTGTTGTCATTTCTTACGCTGAAGCAG 278
Db 206 CGCTCATGCACTTACGACTCGGGGAAATTCCTACCGTCACTTTCTGTCCTCCGGAATGG 265
QY 279 CCAAGAGGTATTGAAAACTCAGGCTGATGCCCTTCGCCCAACGCCCTATCGTCTTGGACG 338
Db 266 CGAAGGAAGTACTAAAACTCAGACCTCGCTTTCGCAACGAGGCCGAACTTGTGGTGG 325
QY 339 CACAGATTGTGTTTTATAATCGGAAAGATGTCTTGTGTTGTTTCATATGGAGATCACTGGA 398
Db 326 CTGACATCGTCCATTATGATAGTACGATATAGCATTTTCTCATATGGTGAATACTGGA 385
QY 399 GGCAGATGAAGAAAATTTGGATACTTGAATTTCTGAGTGCCTCAAAAAGTTCAATCTCTCA 458
Db 386 GGCAGATTCTGTAATTTTGCATCTCGAACTCCTTAGTGCCAAAGATGGTCAAAATTTCTTTA 445
QY 459 GGTTAATCCGAGAGGAAGAAATGGAGGATGCCATCACATTCCTCCCTTCGAAAAGCCGGAT 518
Db 446 GCTCAATTCCGCCAGGATGAGTGTGATGATGCTCATCTATACGAACTATGCCAAAT 505
QY 519 CTCGGTCAATATTACAAAGATCAATTTATGGCATTAATAATTTTCGATCATGATGAAGACAT 578
Db 506 TTCCCGTCAACCTTACAGACAAAATATTTTGGTTTACAAGTTCCGTAATCTTGTAGATCAG 565
QY 579 CCGTTGGTAA---TTGTAAGCAAAAGAAAGATTGCTGAGTGTTCGCGATGCTCAATG 635
Db 566 CTCGCGAAAAATATGCTGTGACCAAGACAAAATGATAATTTTCATGAGGAAATAATAT 625
QY 636 AGGCGACCGACGAGTTTTCGACCGCAGACGCTTTTCCGACGTGGAAATTTACTTCACTATA 695
Db 626 CATTCACAGGTGGAATTTAGTATTGCTGATTTTTCCTTACATGGAATAATGCTACATGATG 685
QY 696 TCATTTGGAGCTAGTCAAAACCCAGCGCTTTCGATCAGAGATGACGATATATCTTTGAAG 755
Db 686 TTGGTGGTTTCAAAAATCTAGACTGCTGAAAGCTCGTAAAAATCGATGAGATTTTGGAAAC 745
QY 756 AGATTCTTAATCAACACAAAGCCAATAAGCCT- - - - -TTTG 791
Db 746 ATGTAGTGAATGACACAAACAGAAATCGAGCGGATGGCCAAAGGGTAATGGCGAATTTG 805
QY 792 AAGCGGATAACTTAATGGATGTTCTATTGAATCTTCAAAAAAATGGAAACGTTTCCAGTGC 851
Db 806 GCGGTGAAGATTGTGATCGATGTTTTCCTAAGGGTTTCGAGAAAGTGGAGAAGTTCAATTT 865
QY 852 CAGTGACAAACGAAAGCATCAAGCATCCGTTTTCGAAATGTTTACTGCCGGAGCGGAAA 911
Db 866 CCATCAGCGATGACAATATCAAAATCAATATATTAGTGACATGTTCTCCGCTGGATCTGAAA 925
QY 912 CAACCTTCGAAAGCTACAGAACTGGGTAAATGGCAGAGCTGATGAAATCCAACTGAACTAA 971
Db 926 CGTATCGACAATTAATTTGGGCAATTAGCTGAATGATGAAGAAACCAAGTGTCTAG 985
QY 972 GAAAAAGCAAGAAGAAGTTAGCAAGTATTTGGTGAATTTGGGAAATGGGAAAAGTTGATGAATCAA 1031
Db 986 CAAAGGCAACAGCTGAAGTGAGACAAGTCTTCAAGGAAAAGAAAGGTTTTTCAACAAATG 1045
QY 1032 GATTTTCATGATTTGAAATTTCTTCAAGTTAGTGGTTTAAAGAAATCTAAGATTTACATCCTC 1091
Db 1046 ATCTTGTAGTGTGAAGTACTTGAAGTTAGTAAATCAAGAAAACTCTTAAGGATGCAACCTC 1105
QY 1092 CGGT- - -TGTCTTGTATTCGAGGGAGTGTAGAGAAACACAGAAATGATGATATGAAA 1148
Db 1106 CAATTCCTCTATTAGTCCCTAGAGAAATGATGAAGGATACAAAGATTTGACGGGTACAATA 1165
QY 1149 TTCTCCGAAACACTCGAATTTGTTGTAATGCTTTGGGCGATAGGAAGATCTCTAATCTT 1208

Db 1166 TACCTTTCAAAACTCGAGTCATAGTAAATGCAATGGGCAATTGGACGAGATCTCGAAAGTT 1225
Qy 1209 GGTCCGAACCTGGAAGTTTAAACCGAGAAAGGTTTAAAGATGTGCAATATGATATAAAG 1268
Db 1226 GGGATGACCTGAAAGCTTTTCCCGAGAGATTCGAGAATAGTCTGTGTGACTTCTTG 1285
Qy 1269 GGACGACATTTGAATGCTGACATTTGGTGGAGGAAAGAAATATGCTCGGCATCTT 1328
Db 1286 GAAGCCATCATCAATTTATTCATTTGGTGGCGGAAGAGGATTTGCTCGGAATGCTTT 1345
Qy 1329 CAGCTATTACCAATTTGGAGTATGTCATTAATAATCTATATATCATTTTAATTCGGAAC 1388
Db 1346 TTGGTTTAGCCAATGTGGACAACCATTAGCTCAATTAATTTATCATTGCAATCGGAAC 1405
Qy 1389 TGGCCGATGGAATTAACCTCAACACATTTGATATGACTGAAGCTATTTGGCGTGTCTCA 1448
Db 1406 TCCCTAATGGACAAAGTCAGAAATTTGGACATGACGGAGTCACCTGGAATTTCTGCA 1465
Qy 1449 GGAATAAATAGATCTTAAGTTGATCTCTATTCATATCAAGTTAGCTTAGGCTCAATA 1508
Db 1466 CAAGAAAGGATGCTGTGTTTGAATGGCACCCCTTATGATCCTTGAATGTATTGAGACA 1525
Qy 1509 TTCTTTGATTACATAGGAGGTTGAAA 1535
Db 1526 GTTGTAGAAATAAAGAGGGAGAAA 1552

RESULT 10
LOCUS CS223966 1673 bp DNA linear PAT 15-DEC-2005
DEFINITION Sequence 222 from Patent WO2005111217.
ACCESSION CS223966
VERSION CS223966.1 GI:83685962
KEYWORDS
SOURCE Nicotiana tabacum (common tobacco)
ORGANISM Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiids; Solanales; Solanaceae; Nicotiana.
REFERENCE 1
AUTHORS Xu,D.
TITLE Nicotiana nucleic acid molecules and uses thereof
JOURNAL Patent: WO 2005111217-A 222 24-NOV-2005;
U.S. Smokeless Tobacco Company (US)
LOCATION/Qualifiers
FEATURES
1..1673
/organism="Nicotiana tabacum"
/mol_type="unassigned DNA"
/db_xref="taxon:4097"
source

ORIGIN
Query Match 23.0%; Score 398.8; DB 2; Length 1673;
Best Local Similarity 57.6%; Pred. No. 1.6e-112;
Matches 809; Conservative 0; Mismatches 562; Indels 33; Gaps 4;
Qy 113 CTTCCAGGCCCATGGAAGTTTCTATCATAGTAAATCTTCTCATTTTATCTCACTTCT 172
Db 135 CCACCAGGTCATGGAACCTACCAATCTAGGAAGTATGCTTCATATGCT-----TGGT 188
Qy 173 GATCTAGGCCCATGAACGTTTATAGAGCTTGGCTCAAATTTATGGACCTGTTATGAGTCTT 232
Db 189 GGACTACCAACCATGTCCTTAGAGATTTAGCCAAAATATAGGACCCTTATGACCTT 248
Qy 233 CAATTTGGCCAAAGTTTACGCTGTGTCAATTTCTCAGCTGAAAGCCAGCAGAGGTTATG 292
Db 249 CAATTAGGTGAAGTTTCTCGGTTGTGGTTACTTCTCTGATACGCGCAAAAGATTA 308
Qy 293 AAAACTCAGGCTGATGCCTTCGCCCAACGCTTATCGTCTTGGACGACAGATGTGTT 352
Db 309 AAAACTCATGACATCGCTTTTGGCTTAGGCCCTTAGCCCTTTTGGCCCCGAGATGTGCTG 368
Qy 353 TATAATCGAAAGATGCTCTGTTTGTCTCATATGGAGATCACTGGAGGCAGATGAAGAAA 412

Db 369 TACAATAGTCTGATCTAGCCCTTTTGGCCCTATGGGCACTATTGGAGACAAATGCGTAAA 428
Qy 413 ATTTGGATPACTTGAATTTCTGAGTGCCAAAAGTTTCAATCTCCAGGTTTAAATCCGAGAG 472
Db 429 ATATGTGTCTTGGAAAGTCTCAGTGCCAAAGATGTTCCGACATTTAGCTCTATTAGCGGG 488
Qy 473 GAAGAAATGGAGGATGCCATCACATTTCTCGTTCGAAAGCGGATCTCCGGTCAATATT 532
Db 489 AATGAAGTTCTTCGCTCTCAATTAATTTTATCCGGTCATCTTCTGGTGAACCTTAATTAAT 548
Qy 533 ACAAAGATCATTTATGGCATTAATAATTCGATCATGATAAGAACATCCGTTGGTAA--T 589
Db 549 ACGAAAGGATCTTTTGTTCACAAGCTCCATGACATGATAGATCAGCGTTTGGGCAAGTG 608
Qy 590 TGTAAAGCAAAAGAAAGATTCCTGAGTGTGCCGATGCGATCAATGAGGCGAGCGAGT 649
Db 609 TTCAAAGAGCAAGACAAATTTATCAACTAATAATAAGAAAGTGATACTCTTAGCAGAGGG 668
Qy 650 TTTGGCACCGCAGAGCGCTTTTCCGACGTGGAAATTTACTTCACTATATCATTTGGAGCTGAG 709
Db 669 TTTGATGGCTGACATATCCCTTCACTGAAGTTTCTTCATGTCTCAGTGGAAATGAAG 728
Qy 710 TCAAAACCCAGCGCTTTTGCATCAGGAGATTGACGATATATCTTTGAAGAGATTTCTTAATGAA 769
Db 729 GGTAAAGATTATGAATGCACACATAGGTAGATGCCATTGTTGAGAAATGTCATCAATGAG 788
Qy 770 CAC-----AAAGCCAAATAGCCCTTTTGAAGCGGATCACTTAATG 808
Db 789 CACAAGAAAAATCTTGCATTTGGGAAAACTAATGGAGCGTTTAGGAGGTGAAGATTTAAAT 848
Qy 809 GATGTTCTATTGAATCTTCAAAAAAATGGAACGTTTCCAGTGCCAGTCACAAAACGAAAGC 868
Db 849 GATGTTCTCTAAGACATTATGAATGATGAGGCGCTTCAATTTCTCATCACCACACGACAC 908
Qy 869 ATCAAAGCATCCGTTTTCGAAATGTTTACTGCGGAGCGAAACAACTTCGAAAGCTACA 928
Db 909 ATCAAAGCTATAATTTTGTGACATGTTTCTGCCGGACGGAGACTTCATCGTCAACAAT 968
Qy 929 GAATGGGTAATGGCAGAGCTGATGAAATCAACTGAACCTAAGAAAAGCACAGAAGAA 988
Db 969 GTGTGGCTATGTTAGAAATGGTGAATAATCCAGCGTATTTCGCGAAAGCTCAAGCAGAA 1028
Qy 989 GTTAGACAAGTATTTGGTGAATGGGAAATGTAATGAATCAAGATTTTCATGATTTGAAA 1048
Db 1029 GTAAAGAGAGCATTTAGAGGAAAGAAACTTTCATGAAATATGATGAGGAGGCTAAAC 1088
Qy 1049 TTCTTTCAAAGTTAGTGTTTAAAGAACTCTAAGATTAATCACTCCGTT---GTCCTGATT 1105
Db 1089 TACCTAAAGTTAGTAATAAAGAACTCTAAGACTTCATCCACCGGTTCCACTTTTGCTC 1148
Qy 1106 CCGAGGAGGTAGAGAAACACACGAATTTGATGATGAATTAATTCATCCGAACTCGA 1165
Db 1149 CCAAGAGAATGTAGGAAAGAGACAAATATAAACCGCTACACTATTTCGTAAAGACCAAA 1208
Qy 1166 ATTTGTGTGAATGCTTGGCGCATAGGAAGAGATCTTAATCTACTTGGTCCGAACTCGAAAG 1225
Db 1209 GTCATGTTAATGTTGGGCTTTGGGAAGAGATCCAAATATTTGGAATGACCGAGAACT 1268
Qy 1226 TTTAAACCAAGAAAGGTTTAAAGATTTGCAATTTGATTAATAAGGACGACATTTGAAC 1285
Db 1269 TTTATGCCAGAGAGATTTGAGCAGTGTCTAAGGATTTTGTGGTAAATAATTTTGAATAT 1328
Qy 1286 GTACATTTGGTGCAGGAAAGAGATATGCTCGCATTTACTTCACTATTTACCAATTTG 1345
Db 1329 CTTCCATTTGGTGCAGGAAAGGAGGATTTGCTCGGATTTTCTTGGCTTAGCTAATGCT 1388
Qy 1346 GAGTATGTCATTATAAATCTATTATATCAATTTTAAATGGGAACCTGGCCGATGGAATTA 1405
Db 1389 TATTTGCCATGGCTCAATTACTATATCACTCGATTGGAACTCCCTGCTGGAATCGAA 1448
Qy 1406 CTTCAAAACATTTGATATGACTGAAGCTATTTGGCGGTGCTCTCAGGAAAAAATATAGAT 1465
Db 1449 CCAAGCGACTTGGACTTGAAGTGGTGGAGTAACTGCGCGCTAGAAAAAGTGAACCTT 1508


```
QY 1466 AAGTTGATTCCTATTCCATATCAA 1489
Db 1509 TACTTGGTTCGGACTCCTTATCAA 1532

RESULT 11
CS223950 1610 bp DNA linear PAT 15-DEC-2005
DEFINITION Sequence 206 from Patent WO2005111217.
ACCESSION CS223950
VERSION CS223950.1 GI:83685854
KEYWORDS
SOURCE Nicotiana tabacum (common tobacco)
ORGANISM Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiales; Solanales; Solanaceae; Nicotiana.
1
REFERENCE
AUTHORS Xu, D.
TITLE Nicotiana nucleic acid molecules and uses thereof
JOURNAL Patent: WO 2005111217-A 206 24-NOV-2005;
U.S. Smokeless Tobacco Company (US)
FEATURES
source Location/Qualifiers
1..1610
/organism="Nicotiana tabacum"
/mol_type="unassigned DNA"
/db_xref="taxon:4097"

ORIGIN
Query Match 22.7%; Score 393.2; DB 2; Length 1610;
Best Local Similarity 56.7%; Pred. No. 8.9e-111;
Matches 825; Conservative 0; Mismatches 598; Indels 33; Gaps 4;

QY 101 AAACAGAAATCCACCTCCAGGCGCATGGAAGTTTCCTATCATAGGTAAATCTTCCTCATTTA 160
Db 106 AAACAAATTTGCTCCTCGTCCATGGAATTAATCACTTTTATTGGAGTTTACACCAATTTG 165

QY 161 TTACTACATCTCTGATCTAGGCGCATGAACGTTTTAGAGCCTTCGCTCAAAATTTATGACCT 220
Db 166 GCTGTGGCAGGTCCACTTCTCACATGGCCTTAAATAATTTAGCCAAAGCTATGTCTCT 225

QY 221 GTTATGAGCTTCAAAATTTGGCCAAAGTTTCAGCTGTTGTCAATTTCTTCAGCTGAAGCAGCC 280
Db 226 CTTATGCAATTTCAACTCTGGCAAAATTTCTACACTCGTCAATCATCATCACTCAAAATGSCA 285

QY 281 AAAGAGTTATGAAGAACTCAGGCTGATGCTTCGCGCCAAAGCCCTATCGTCTTTGACGCA 340
Db 286 AAAGAAGTACTAAAAAACTCAGCACTCGCTTTTGGCCACTAGACCAAAAGCTTGTCTGTGGCC 345

QY 341 CAGATTGTGTTTTATAATCGGAAAGATGCTTTGTTGCTTCATATGGAGATCACTGGAGG 400
Db 346 GACATCATCTACGACAGCAGGACATAGACATTTCCGCATACGGTGAATACTGGAGA 405

QY 401 CAGATGAAGAAATTTGGATCTTGAATTTCTGAGTGGCCAAAAAGTTTCAATCTCCAGG 460
Db 406 CAAATTCGTAATTTGTCATATTGGAATCTTTGAGTGGCCAAAGTGTCAAGTTTTTTAGC 465

QY 461 TTAATCCGAGAGAAATAGAGATGCCATCAATTCCTCGTTTCGAAAGCCGATCT 520
Db 466 TCGATTCCGCAAGATGAGCTCTCGAAGATGGTTTCATCTATACGAACGACGCCCAATCTT 525

QY 521 CCGGTCAATATTACAAAGATCATTTATGGCATTATAATTTTCGATCATGATAGAACATCC 580
Db 526 CAGTCATCTTACGACAGAGATTTTGGTTTACGAGTTCGGTAAATTTGTAGATCAGCT 585

QY 581 GTTGTAA---TTGTAAGCAAAAGAAAGATGCTGAGTGTTCGCGATGCAATGAG 637
Db 586 TTAGGGAAGATATGTGGTGACCAAGACAAATTTGATCATTTTTATGAGGGAATAATATCA 645

QY 638 GCAGCGACGAGTTTGGCACCCGACGCTTTTCCGAGTGGAAATTTACTTCATATATC 697
Db 646 TTGGCAGGTGGATTTAGTATTCGTGATTTTTCCTACATGGAATATGATTCATGATAT 705

698 ATTGGAGCTGAGTCAAAAACCAGGCGTTTGCATCAGGAGATTGCAGATATATCTTGAAG 757
706 GATGGTTCAAAATCTAAACTGGTGAAGGCACATCTGTAAGATTTGATGAAATTTTGGAAAT 765
758 ATTTCTTAATGAACACAAA-----GCCAATAAGCCTTTTGAA 793
766 GTGGTAAATGAGCACAAACAGAAATCGAGCAGATGGTAAAAAGGGTAATGGTGAATTTGGT 825
794 GCGGATACTTAATGGATGTTCTTAATTGAATCTTTCAAAAAAATGGAACGTTTCAGTGGCCA 853
826 GGAGAAGATCTGATTTGTTTAAAGAGTTAGAGAAAGTGGAGAAGTTCAAAATTTCCA 885
854 GTGACAAACGAAAGCATCAAAACATCCGTTTTCGCAAAATCTTACTCGCGGAGCGGAACA 913
886 ATCACAGATGACAAATATCAATCAATTAATCGACATGTTCTCTCCGATCGGAACA 945
914 ACTTCGAAAGCTACAGAATGGGTAAATGGCAGAGCTGATGAAAAAATCAAACTGAACATAAGA 973
946 TCATCGACAACTATAATTTGGGCAATTAGCTGAAATGATGAAGAACCAAGTGTTTTAGCA 1005
974 AAAGCACAGAAGAGTTAGACAAAGTATT---TGGTGAATGGGAAAAAGTTGATGAATCA 1030
1006 AAAGGCACAGCTGAAGTGAGCCAAGCTTTGAAGGGGGAAGAAAAATTAGTTTCAAGAGATT 1065
1031 AGATTTTCATGATTGAAAATTTCTTCAAGTTAGTGGTTTAAAGAACTCTTAAGATTTACATCCT 1090
1066 GATATTGATTAAGCTAAAGTATTGAGTTAGTATCAAAAGAACTTTTAAAGATTCACACCT 1125
1091 CCGGT---TGTCTTGATTCGAGGGAGTGTAGAGAAACAAACAGAAATTTGATGATATGA 1147
1126 CCAATTCCTCTGTAGTCCCTAGAGAAATGATGGAAGATACAAAGATTGATGGTTACAAT 1185
1148 ATTCAATCGAACACTCGAATTTGTTGTAATGCTTGGGCGCATAGGAAGAGATCCTAATACT 1207
1186 ATACCTTTCAAAACAAGAGTCAATTTGTTAATGCTATGGCAATTTGGACGAGATCCTCAAAGT 1245
1208 TGTGCGGAACCTGGAAGTTTAAACCCAGAAAGTTTAAAGATTGTCGAATTTGATTTATA 1267
1246 TGGGATGATCTCGAAAGCTTTACGCGAGAGATTTGAGAATAATTTCTATTGATTTCTT 1305
1268 GGGACGACATTTGAACTGATACCTATTTGTCAGGAAAAAGAAATATGCTCGCATTACT 1327
1306 GGAATATCATCATCAATTTATTCATTTGTCGAGGAAGAGATTGTCCTGGAATGCTA 1365
1328 TCAGCTATTACCAATTTGGAGTATGTCATTATAAATCTATTATATATCATTTTAAATTTGGAA 1387
1366 TTTGTTTAGCTAATTTGGACAACCTTTTAGCTCAGTTACTTTATCACTTCGATTGGAAA 1425
1388 CTGGCCGATGGAATTTACACCTCAAACTTTGATATGACTTGAAGCTATTTCGCGGTGCTCTC 1447
1426 CTCCTAATGGAACAACTACCAAAATTTTCGACATGACTGACTGACCTGGAAATTTCTGCT 1485
1448 AGGAAAAAATAGATCTTAAAGTTGATTCCTATTTCATATCAAGTTAGCTTAGGCTCAAAAT 1507
1486 ACAAGAAAGGATGATCTTATTGTTGATGTCACCTCTGCTCATTTCTTGATTAGTATGCT 1545
1508 ATTTCTTGATACATA 1523
1546 GCTTTTCTATTGGAGA 1561

RESULT 12
CS224014
LOCUS CS224014
DEFINITION Sequence 270 from Patent WO2005111217.
ACCESSION CS224014
VERSION CS224014.1 GI:83685886
KEYWORDS
SOURCE Nicotiana tabacum (common tobacco)
ORGANISM Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
```

asterids; lamids; Solanales; Solanaceae; Nicotiana.

REFERENCE

AUTHORS	Xu, D.
TITLE	Nicotiana nucleic acid molecules and uses thereof
JOURNAL	Patent: WO 200511217-A 270 24-NOV-2005;

U.S. Smokeless Tobacco Company (US)

```

FEATURES
    source
        Location/Qualifiers
            1..1610
                /organism="Nicotiana tabacum"
                /mol_type="unassigned DNA"
                /db_xref="taxon:4097"

```

ORIGIN

Query Match 22.7%; Score 393.2; DB 2; Length 1610;
Best Local Similarity 56.7%; Pred. No. 8.9e-111;
Matches 825; Conservative 0; Mismatches 598; Indels 33;

Qy	101	AAACAGAAATCCACCTCCAGGGCCATGGAAGTTTCTCTATCATAGAGTAATCTTCTCCTCATTTA	160
Db	106	AAACTAAATTTGGCTTCCTGGTCCATGGAAATTTACCTTTTATTGGAAAGTTTACACCATTTG	165
Qy	161	TTACTCACCTTCGTACTAGGCCATGAAACGTTTTAGAGCCTTGGCTCAAAATTTATGACCT	220
Db	166	GCTGTGGCAGGTCCACTTCTCTCACATGGCCTAAAAAATTTAGCCAAACGCTATGGTCTT	225
Qy	221	GTTATGAGTCTTTCAAATTTGGCCAAAGTTTTCAGCTGTGTTCATTTCTCAGCTGAAACGACC	280
Db	226	CTTATGCACTTTACAACTTTGGACAAATTCCTACACTCGTCATATCATCACTACCTCAAAATGGCA	285
Qy	281	AAAGAGGTTATGAAAACCTCAGGCTGATGCCITCGGCCCAACGCCCTATCGTCTTGACGCA	340
Db	286	AAAGAAGTACTAAAAAATCAAGACCTCGCTTTTGCCACTAGAACCAAGCTTGTCTGGGCC	345
Qy	341	CAGATTGTGTTTTATAATCGGAAAGATGCTTTGTTTGTCTTATATGGAGATCACTGGAGG	400
Db	346	GACATCATTTCACTACGACAGCAGGACATAGACATTTCCGCATACGGTGAATACTGGA	405
Qy	401	CAGATGAAGAAATTTTGGATTACTTGAAATTTCTGAGTGGCCAAAAGTTCAATCTCTCCAGG	460
Db	406	CAAAATTCGTAAATTTTGCAATATTGGAATCTTTGAGTGGCCAAAGTGGTCAAGTTTTTTTAGC	465
Qy	461	TTAATCCGAGAGGAAGAAATGGAGATGCCATCACTTCCGTTTCGAAAGCCGGATCT	520
Db	466	TCGATTCCGCAAGATGAGCTCTCGAAGATGGTTTCATCTATACGAACGACGCCCAATCTT	525
Qy	521	CCGGTCAATATTACAAAGATCAATTTATGGCATTTATAATTTTCGATCATGATAAGAACATCC	580
Db	526	CCAGTCAATCTTTACCGACAAGATTTTTTGGTTTTACGAGTTTCGGTAATTTGTAGATCAGCT	585
Qy	581	GTGTGTAA---TTGTGAAGCAAAAGAAAGATTGCTGAGTGTGGCGATGCAGTCAATGAG	637
Db	586	TTAGGGAAGATATGTGGTGCCAAAGACAAATTTGATCATTTTTATGAGGGAATAATATCA	645
Qy	638	GCAGCGCAGGTTTTTGGCACCCGACAGCGTTTTCCGACGTGGAAATTACTTCACCTATATC	697
Db	646	TTGGCAGGTGGAAATTTAGTATTGCTGATTTTTTCCCTACATGGAAGAAATGATTCATGATTT	705
Qy	698	ATTGGAGCTGAGTCAAAACCCAGCGGCTTTGTCATCAGGAGATTGACGATATATCTTGAAGAG	757
Db	706	GATGCTTCAAAATCTAAACTGCTGGAAGGCACATCGTAAGATTTGATGAAATTTTGGAAAAAT	765
Qy	758	ATTCCTTAATGAACACAAA-----GCCAATTAAGCCTTTTGGAA	793
Db	766	GTGGTAAATTGACGACAAAACAGAAATCGAGCAGATGTTAAAGGGTAATGGTGAATTTGGT	825
Qy	794	GCGGATAACTTAATGATCTTCTATTGAAATCTTCAAAAAAATGGAAAACTTCACAGTGCCA	853
Db	826	GGAGAAGATCTGATTTGATGTTTTTGTTTAAGAGTTAGAAAAGTTGGAGAAGTTCAAAATCCA	885
Qy	854	GTGACAAAACGAAGCATCAAGCATCCGTTTTTGCAAATGTTTTTACTCCCGGGAGCGAAACA	913
Db	886	ATCAGAGATGACAAATATCAAAATCAATATTAATTCGATCTTCTCTGCCGGATCGGAACA	945

Qy	914	ACTTCCGAAGCTACAGNATGGGTAAATGGCGAGCTCGATGAAAAATCCAACTGAACCTAAGA	973
Db	946	TCATTCGCAACTATATTTTGGGCATTAGCTGGAATGATGAAGAACCAAGCTGTTTAGCA	1005
Qy	974	AAAGCAACAAGAAGAGTTAGACAAGTATT--TGGTGAATGGGAAAGTTGATGAATCA	1030
Db	1006	AAGSCAACAGCTGAAGTGAAGCCAAAGCTTTGAAGGGGAAGAAAATTAGTTTTCAGAGATT	1065
Qy	1031	AGATTTTCATCATTTTGAATTTCTTCAAGTTAGTGGTTTAAAGAAAACCTCTAAGATTACATCCT	1090
Db	1066	GATATTCATTAAGCTAAGAATATTTGAAGTTAGTGCATCAAGAGAACTTTTGAAGATGCACCT	1125
Qy	1091	CCGGT---TGCTTTGATTCGAGGGAGTGTAGAGAAACAACGAATTTGATGGATATGAA	1147
Db	1126	CCAAATTCCTCTGTTAGTCCCTAGAGAATGTATGGAAGATACAAAGATTGATGGTTTACAAT	1185
Qy	1148	ATTCAATCCGAACACTCGAAATGTTGTGTAATGCTTTGGCCGATAGGAAGAGATCCTAATACT	1207
Db	1186	ATACCTTTCAAACAAGAGTCAATGTTAATGCATGGCAATTCGACGAGATCCTCAAAGT	1245
Qy	1208	TGTFCGGAACCTCGAAAGTTTAAACCCAGAAAGGTTTAAAGATTGTGCAATTGATTTATAA	1267
Db	1246	TGGGATGATCCTGAAGCTTTAGCCAGAGAGAAATTTGAGAATAATCTATTGAATTTCTT	1305
Qy	1268	GGGACGACATTTGAACCTGGTACCATTTTGGTGCAGGAAAAGAAATATGTCTCGCATTTACT	1327
Db	1306	GGAAATCATCATCTAAATTTTATTCATTTTGGTGCAGGAAGAGGATTTGTCTCGAAATGCTA	1365
Qy	1328	TCAGCTATTACCAATTTTGGAGTAGTGCATTTATAAATCTATTATATCAATTTTAATTCGGAA	1387
Db	1366	TTTGGTTTAGCTAATGTTTGGACAACCTTTAGCTCAGTTACTTTATCACATTCGATTTGGAA	1425
Qy	1388	CTGGCCGATGGAATTTACCTCTCAAAACACTTTTGATATGACTGAAGCTATTTGCGGTGCTCTC	1447
Db	1426	CTCCCTAATGGACAAACTCACCAAAAATTTTCGACATGACTGAGTCACCTGGAAATTTCTGCT	1485
Qy	1448	AGGAAAAAATAGNCTTTAGTTGATTCCTTATTCCTATATCAAGTTAGCTTAGCGTCAAAT	1507
Db	1486	ACAAGAAGGATGATCTATTTTGAATGGCCACCTCTGCTCATCTCTTGATTTAAGTATTGCT	1545
Qy	1508	ATTTCTTTGATTACATA	1523
Db	1546	GCATTTCTATTGGAGA	1561

DECEMBER 1972

RESULTS	LOCUS	SEQUENCE	DEFINITION	CS223944	1566 bp	DNA	linear	PAT 15-DEC-2005
CS223944	LOCUS	Sequence	200 from Patent WO2005111217.	CS223944	1566 bp	DNA	linear	PAT 15-DEC-2005

ACCESSION	ITERATION	CT-82605051
CSZ23944	CSZ23944	1

VERSION
VER:00000
CSZ23244.I GT:03003031

KEYWORDS
cointegration
tobacco
tobacco (comm)

SOURCE
NICOTIANA TABACUM (COMMON TOBACCO)
NICOTIANA TABACUM

ORGANISM *Nicotiana glauca*
Eukaryota: Viridiplantae: Streptocarpales: Solanaceae: *Nicotiana glauca*

Eukaryota; Viridiplantae; Streptophyta; Charophyta; Magnoliophyta; Eudicotyledones; Malvales; Malvaceae; *Albizia*

agasterids: Lamiids: Solanales: Solanaceae; Magnoliophyta; eudicotyledons

REFERENCE 1

REFERENCE
AUTHORS
XIII D.

TTTt.E Nicotiana nucleic acid molecules a

JOURNAL
Patent: WO 200511217-A 200 24-NOV

U.S. Smokeless Tobacco Company (US

FEATURES	Location/Qualifiers
1. <i>the</i>	the
2. <i>the</i>	the
3. <i>the</i>	the
4. <i>the</i>	the
5. <i>the</i>	the
6. <i>the</i>	the
7. <i>the</i>	the
8. <i>the</i>	the
9. <i>the</i>	the
10. <i>the</i>	the
11. <i>the</i>	the
12. <i>the</i>	the
13. <i>the</i>	the
14. <i>the</i>	the
15. <i>the</i>	the
16. <i>the</i>	the
17. <i>the</i>	the
18. <i>the</i>	the
19. <i>the</i>	the
20. <i>the</i>	the
21. <i>the</i>	the
22. <i>the</i>	the
23. <i>the</i>	the
24. <i>the</i>	the
25. <i>the</i>	the
26. <i>the</i>	the
27. <i>the</i>	the
28. <i>the</i>	the
29. <i>the</i>	the
30. <i>the</i>	the
31. <i>the</i>	the
32. <i>the</i>	the
33. <i>the</i>	the
34. <i>the</i>	the
35. <i>the</i>	the
36. <i>the</i>	the
37. <i>the</i>	the
38. <i>the</i>	the
39. <i>the</i>	the
40. <i>the</i>	the
41. <i>the</i>	the
42. <i>the</i>	the
43. <i>the</i>	the
44. <i>the</i>	the
45. <i>the</i>	the
46. <i>the</i>	the
47. <i>the</i>	the
48. <i>the</i>	the
49. <i>the</i>	the
50. <i>the</i>	the
51. <i>the</i>	the
52. <i>the</i>	the
53. <i>the</i>	the
54. <i>the</i>	the
55. <i>the</i>	the
56. <i>the</i>	the
57. <i>the</i>	the
58. <i>the</i>	the
59. <i>the</i>	the
60. <i>the</i>	the
61. <i>the</i>	the
62. <i>the</i>	the
63. <i>the</i>	the
64. <i>the</i>	the
65. <i>the</i>	the
66. <i>the</i>	the
67. <i>the</i>	the
68. <i>the</i>	the
69. <i>the</i>	the
70. <i>the</i>	the
71. <i>the</i>	the
72. <i>the</i>	the
73. <i>the</i>	the
74. <i>the</i>	the
75. <i>the</i>	the
76. <i>the</i>	the
77. <i>the</i>	the
78. <i>the</i>	the
79. <i>the</i>	the
80. <i>the</i>	the
81. <i>the</i>	the
82. <i>the</i>	the
83. <i>the</i>	the
84. <i>the</i>	the
85. <i>the</i>	the
86. <i>the</i>	the
87. <i>the</i>	the
88. <i>the</i>	the
89. <i>the</i>	the
90. <i>the</i>	the
91. <i>the</i>	the
92. <i>the</i>	the
93. <i>the</i>	the
94. <i>the</i>	the
95. <i>the</i>	the
96. <i>the</i>	the
97. <i>the</i>	the
98. <i>the</i>	the
99. <i>the</i>	the
100. <i>the</i>	the

source
1. .1566

/organism="Nicotiana tabacum"

```
/mol_type="unassigned DNA"
```

/db_xref="taxon:4097"

ORIGIN

Query Match	Score	Score 388.8
22.4%	22.4%	22.4%

Best Local Similarity 56.5%; Pred. No. 2

Matches 801; Conservative 0; Mismatch 0; Mismatch 0

QY 101 AAACAGAAATCCACTCCAGGCCCATGGAAGTTTCTATCATAGTAATCTTCTCATTTA 160
DB 98 AAACATAATTTGGCTCTGCTCCATGGAATTAACCTTTTATTGGAAGTTTACACCAATTG 157
QY 161 TTACTCATTCTGATCTAGGCCATGAAGCTTTTAGAGCCTTGCTCAAAATTTATGACCT 220
DB 158 GCTGTGGCAGGTCACCTTCTCACCATGGCCTTAATAAAATTTAGCCAAACGCTATGGTCT 217
QY 221 GTTATGAGTCTTCAAAATTTGGCCAAAGTTTCAAGCTGTTGTCAATTTCTTTCAGCTGAAGCAGC 280
DB 218 CTTATGCAATTTACAATTTGGACAAATTTCTACACTCATCATATCATCACTCAATATGGCA 277
QY 281 AAAGAGGTATGAATACTCAGGCTGATGCTTCGCCCAACGCCCTATCGTCTTTGGACGCA 340
DB 278 AAAGAGGTATGAATACTCAGGCTGATGCTTCGCCCAACGCCCTATCGTCTTTGGACGCA 337
QY 341 CAGATTGCTTTTATTAATCGGAAGATGCTTTGTTGCTTTCATATGAGATCACTGGAGG 400
DB 338 GACATCAATTCACCTACGACAGCAGGACATAGCATTTTCTCCGTACGGTGAATCTGGAGA 397
QY 401 CAGATGAAGAAATTTGGGATCTTGAATTTCTGAGTGCCAAAGAAAGTTTCAATCTCCAGG 460
DB 398 CAAATTCGTAATTTGCAATTTGGAACCTTGGAGTCCCAAGATGGTCAAAATTTTATG 457
QY 461 TTAATCCGAGAGGAAGAAATGGAGGATGCCATCACATTTCTCCGTTTCGAAAGCCGATCT 520
DB 458 TCGATTTCGCCAAGATGAGCTCTCGAAGATGCTCTCATCTATACGAGGACACCCAATCTT 517
QY 521 CGGTCATATTTACAAGATCAATTTATGGCATTATTAATTTTCATCATGATGAAGACATCC 580
DB 518 ACAGTCAATCTTACTGACAAAAATTTTGTGTTTACGAGTTCGGTAACTTGTAGATCAGCT 577
QY 581 GTTGGTAA--TTGTAAGCAAAAGAAAGATGCTGAGTGTGCCGATCGATCAATGAG 637
DB 578 TTAGGAAGATATGTTGACCAAGCAAAATTTGATCAATTTTATGAGGAATAATATCA 637
QY 638 CGACGACGAGTTTGGCACCGCAGACGCTTTTCGACGTGGAAATTTACTTCACTATATC 697
DB 638 TTGGCAGGTGGATTAGTATTCTGATTTTTCCTTACATGGAATGATTCATGATAT 697
QY 698 ATTGAGCTGATCAAAACCCAGGCTTGTGATCAGGAGATTGACGATATCTTGAAGAG 757
DB 698 GATGTTTGAATCTAAATCTGGTGAAGCACATCGTAAGATTGATGAATTTTGGGAAT 757
QY 758 ATCTTAAATGAACACAAA-----GCCAATAAGCTTTTGAA 793
DB 758 GTTGTGATGACACAAAAGAACAGACGAGATGGCAAGAGGTAAATGGTGAATTTGT 817
QY 794 GCGGATACTTAATGATGTTCTATTGAATCTTCAAAAAATTTGAAACGTTCCAGTGCCA 853
DB 818 GGTGAAGATTTCATTTGATGATATTGTTAAGAGTTAGAGAAAGTGGAGNAGTTCAAATTCCT 877
QY 854 GTGACAAAGCAAGCATCAAGCATCCGTTTGTGAAAATGTTTACCTGCGGGAGCGAAACA 913
DB 878 ATCAAAATGCAAAATCAATCAATATTAATTCGACATGTTCTCTCGGGGATCTGAAACA 937
QY 914 ACTTCGAAAGCTACAGATGGGTAAATGGCAGAGCTGATGAAAAATCCAACTGAATCAAGA 973
DB 938 TCATCGACGACTATTAATTTGGGCATAGCTGAATATGATGAAGAACCAAGTGTGTTAGCA 997
QY 974 AAAGCAAGAGAAAGTTAGCAAGTATTTGGTGAATTTGGGAAAAAGTTGATGAATCAAGA 1033
DB 998 NAGGCACAAGCTGAAGTAAGGCAAGCTTTGAAGGAGAAAAAGGTTTCAACAGATTGAT 1057
QY 1034 TTTCAATGATTTGAAATTTCTCAAGTTAGTGGTTAAAGAAACTCTAAGATTATCATCTCGG 1093
DB 1058 CTTGATGAGCTAAAAATATCTCAAGTTAGTAAATCAAAAGAAACCTTAAAGAAATGCAACCTCCA 1117
QY 1094 GT---TGCTTCTGATTCGAGGAGGTAGAGAAACAACACGAATTTGATGGATATGABATT 1150
DB 1118 ATTCCTCTATTAGTTCTTAGAATGTTATGGAGGATACAAGATTGATGGTTTACAATATA 1177
QY 1151 CATCCGAACACTCGAATTTGTTGTAATGCTTGGGCGATAGGAAGAGATCTTAATACTTTGG 1210

DB 1178 CCTTTCAAAACAGAGTCATAGTTAATGATGGCAATCGGACGATCCAGAAAAGTTGG 1237
QY 1211 TGGAACTTGGAAGAGTTTAAACCCAGAAAGTTTAAAGATTGTGCAATTTGATTATAAGGG 1270
DB 1238 GATGACCCCGAAAGCTTTATGCCAGAGAGATTTGAGATAGTTCTATTGACTTTCTTGA 1297
QY 1271 ACAGACATTTGAACTGCTACCATTTGTCGAGGAAAGAAATATGCTCGCATTAATTCTCA 1330
DB 1298 AATCATCATCAGTTTATACATTTGGTCGAGGAAGAGATTGTCCGGAAATGCTATTT 1357
QY 1331 GCTATTACCAATTTGGAGTATGTCATATATAATCTATTATATCATTTTAAATTCGGAACTG 1390
DB 1358 GGTTTAGCTAATTTGGACAACTTTAGCTAGTTACTTATCACTTCGATTTGGAACCTC 1417
QY 1391 GCGGATGGAATTTACACTCAACACTTTGATATGATGATGAGCTGATTCGGCGGTCTCTCAGG 1450
DB 1418 CCTAATGGACAAAGTCATGAGAAATTTGACATGACTGACTGACCTGGAATTTCTGCTACA 1477
QY 1451 AAAAAAATAGATCTTAAAGTTGATTCCTATTCCATATCA 1488
DB 1478 AGAAGGATGATCTTGTGTTGATTCCTCTCTATGA 1515

RESULT 14
CS223902
LOCUS CS223902 1576 bp DNA linear PAT 15-DEC-2005
DEFINITION Sequence 158 from Patent WO2005111217.
ACCESSION CS223902
VERSION CS223902.1 GI:83685432
KEYWORDS
SOURCE Nicotiana tabacum (common tobacco)
ORGANISM Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiids; Solanales; Solanaceae; Nicotiana.

REFERENCE 1
AUTHORS Xu D.
TITLE Nicotiana nucleic acid molecules and uses thereof
JOURNAL Patent: WO 200511217-A 158 24-NOV-2005;
U.S. Smokeless Tobacco Company (US)
LOCATION/Qualifiers
FEATURES
source 1..1576
/organism="Nicotiana tabacum"
/mol_type="unassigned DNA"
/db_xref="caxon:4097"

ORIGIN

Query Match 22.4%; Score 388.8; DB 2; Length 1576;
Best Local Similarity 56.5%; Pred. No. 2.1e-109;
Matches 801; Conservative 0; Mismatches 587; Indels 30; Gaps 3;

QY 101 AAACAGAAATCCACTCCAGGCCCATGGAAGTTTCTATCATAGTAATCTTCTCATTTA 160
DB 110 AAACATAATTTGGCTCTGCTCCATGGAATTAACCTTTTATTGGAAGTTTACACCAATTG 169
QY 161 TTACTCATTCTGATCTAGGCCATGAAGCTTTTAGAGCCTTGCTCAAAATTTATGACCT 220
DB 170 GCTGTGGCAGGTCACCTTCTCACCATGGCCTTAATAAAATTTAGCCAAACGCTATGGTCT 229
QY 221 GTTATGAGTCTTCAAAATTTGGCCAAAGTTTCAAGCTGTTGTCAATTTCTTTCAGCTGAAGCAGC 280
DB 230 CTTATGCAATTTACAATTTGGACAAATTTCTACACTCATATCATCACTCAATATGGCA 289
QY 281 AAAGAGGTATGAATACTCAGGCTGATGCTTCGCCCAACGCCCTATCGTCTTGGACGCA 340
DB 290 AAAGAGGTATGAATACTCAGGCTGATGCTTCGCCCAACGCCCTATCGTCTTGGACGCA 349
QY 341 CAGATTGCTTTTATTAATCGGAAGATGCTTTGTTGCTTTCATATGAGGATCACTGGAGG 400
DB 350 GACATCAATTCACCTACGACAGCAGCAGCATAGCATTTTCTCCGTACGGTGAATACTGGA 409
QY 401 CAGATGAAGAAAAATTTGGGATCTTGAATTTCTGAGTGCCAAAAAGTTTCAATCTCTCCAGG 460

Db 410 CAAATTCGTAATAATTTGCATATTTGGAACCTCTTGAGTGCACAAGATGSGTCAAAATTTTTTATG 469
QY 461 TTAATCCGAGAGGAAGAATGAGGATGCGCATACATCTCTCCGTTTCGAAAGCCGATCT 520
Db 470 TCGATTCGCCAAGATGAGCTCTCGAAGATGCTCTCATCTATACGAACGACACCCCAATCTT 529
QY 521 CGGTCATATTTACAAAGATCAATTTATGGCAATTAATAATTTCCGATCATGATGAACATCC 580
Db 530 ACAGTCAATCTTACTGACAAAATTTTTTGGTTTACAGTTCGGTAACTGTGTAGATCAGCT 589
QY 581 GTTGATAA---TTGTAAGCAAAAAGAAATGCTGAGTGTGCGATGCGAGTCAATGAG 637
Db 590 TTAGGGAAGATATGTTGGTACCAAGCAAAATTTGATCATTTTTATGAGGGAATAATATCA 649
QY 638 GCAGCGAGGATTTTGGCACCGCAGACGCTTTTCGAGCTGGAATTTACTTCACATATC 697
Db 650 TTGGCAGGTGGATTTAGTATGCTGATTTTTTCCCTACATGGAATAATGATCATGATAT 709
QY 698 ATTTGAGCTGAGTCAAAACCCAGGCGTTTGCATCAGGAGATTCAGCATATACTTGAAGAG 757
Db 710 GATGCTCGAATCTAACTGGTGAAGCACATCGTAAGATTTGATGAATTTTGGCAAT 769
QY 758 ATTCCTAATGAACAAA-----GCGAATAAGCCTTTTGA 793
Db 770 GTTGTGATGACACAAAAGAACAGACAGATGCGAAGGGTAAATGGTGAATTTGGT 829
QY 794 GGGGATAACTTAATGGATGTTCTATTGAATCTTCAAAAATATGGAACGTTCCAGTGCCA 853
Db 830 GGTGAAGATTTGATTTGATGTTTGAAGATTTAGAGTTAGAGAAAGTGGAGAGTTCAATTCCT 889
QY 854 GTGCAAAACGAAAGCATCAAGCATCCGTTTTCGAAATGTTTACTGCGGGAGCGAAACA 913
Db 890 ATCACAATGACAATATCAATCAATATTAATCGACATGTTCTCTGCAAGGATCTGAAACA 949
QY 914 ACTTCGAAGCTACAGAATGGTGAATGCGAGAGTGAAGAAATCCAACTGAATGA 973
Db 950 TCATCGACGACTATAATTTGGGCAATTAATCGACATGTTCTCTGCAAGGATCTGAAACA 1009
QY 974 AAAGACAAGAAGAGTTAGACAAGTATTTGTGAAATGGGAAAAGTTGATGAATCAAGA 1033
Db 1010 AAGGCACAAGCTGAATGAGGCAAGCTTTGAAGGAGAAAGGTTTCAACAGATTGAT 1069
QY 1034 TTTCAATGATTTGAAATCTTCAAGTTAGTGGTTTAAAGAACTCTTAAGATTTACATCCTCG 1093
Db 1070 CTTGATGAGCTAAAATATCTCAAGTTAGTAAATCAAGAAACCTTAAGAATGCAACCCTCCA 1129
QY 1094 GT---TGCTCGATTCGAGGAGTGTAGAGAAACAACACGAATTTGATGCGATGAAAT 1150
Db 1130 ATTCCTCTATTAGTTCCTAGAGAAATGATGAGGATACAAAGTTGATGTTTACAATATA 1189
QY 1151 CATCGAACACCTCGAATTTGCTGAATGCTTGGCGATAGGAAGAGATCTTAATCTTGG 1210
Db 1190 CTTTTCNAAACAAGAGTCATAGTTAATGATCGGCAATCGGACGAGATCCAGAAAGTTGG 1249
QY 1211 TCGGAACCTGGAAGTTTAAACCCAGAAAGTTTAAAGATTTGCAATGATTAATAAGGG 1270
Db 1250 GATGACCCCGAAAGCTTTATGCCAGAGATTTGAGAAATGATTTGACTTTCTTGA 1309
QY 1271 ACACATTTGAATGGTACCATTTGGTGCAGGAAAGAAATATGCTCGGCAATTAATCA 1330
Db 1310 AATCATCATCAGTTTATACATTTGGTGCAGGAAGAGATTTGCGGGAATGCTATTT 1369
QY 1331 GCTATTATCAAAATTTGGAGTATGCTCAATTAATAATCTATTATATCAATTTAAATTTGGAAC 1390
Db 1370 GGTTTAGCTAATTTGGACAACCTTTAGCTCAGTTACTTTATCACTTCGATTTGGAACCTC 1429
QY 1391 GCGGATGGAATTTACACCTCAACACTTTGATATGATGAGAGCTATTTGGCGGTCTCTCAGG 1450
Db 1430 CCTAATGGCAAAAGTCATGAGAATTTTCGACATGACTGAGTCACTCGGAATTTCTGCTACA 1489
QY 1451 AAAAAAATGATCTTAAGTTGATTTCTTATTTCCATATCA 1488

Db 1490 AGAAAGGATGATCTTGTGTTTGAATGGCACTCCTTATGA 1527

RESULT 15
CS223904
LOCUS CS223904 1581 bp DNA linear PAT 15-DEC-2005
DEFINITION Sequence 160 from Patent WO2005111217.
ACCESSION CS223904
VERSION CS223904.1 GI:83685433
KEYWORDS
SOURCE Nicotiana tabacum (common tobacco)
ORGANISM Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Nicotiana.

1
Xu,D.
Nicotiana nucleic acid molecules and uses thereof
Patent: WO 2005111217-A 160 24-NOV-2005;
U.S. Smokeless Tobacco Company (US)
Location/Qualifiers
1. .1581
/organism="Nicotiana tabacum"
/mol_type="unassigned DNA"
/db_xref="taxon:4097"

FEATURES
source

ORIGIN

Query Match 22.4%; Score 398.8; DB 2; Length 1581;
Best Local Similarity 56.5%; Pred. No. 2.1e-109;
Matches 801; Conservative 0; Mismatches 587; Indels 30; Gaps 3;

QY 101 AAACAGAAATCCACCTCCAGGCCCATGGAGTTTCTCTATCATAGGTAATCTTCCTCATTTA 160
Db 115 AAATCAAAATTTGCGCTCTGCTGTCATGGAAATTAACCTTTTATTTGGAAGTTTACACCATTTG 174
QY 161 TTACTCACTTCTGATCTAGGCCATGAACGTTTTTAGAGCTTTGGCTCAAAATTTATGAGACT 220
Db 175 GCTGTGGCAGGTCCACTTCCTCACCAATGGCTTAAATAATTTAGCCAAAGCTATGGTCT 234
QY 221 GTTATGAGTCTTCAAAATTTGGCCAAAGTTTCAAGCTGTGTGTCATTTCTTCACTGGAAGCAGCC 280
Db 235 CTTATGCAATTTACAACTTTGGACAAATCTCACTACATCATATCATCACTCAAAATGGCA 294
QY 281 AAAGAGTTATGAACCTCAGGCTGATGCTTCGCCCAACGCCCTATCGTCTTGGACGCA 340
Db 295 AAAGAAGTACTAAAACTCACGACCTCGCTTTTGCCACTAGACCAAAAGCTTGTCTGTGGCC 354
QY 341 CAGATTGTGTTTTATAATCGGAAGATGCTTTGTTTGTCTTATATGAGATCACTTGGAGG 400
Db 355 GACATCATCTACGACGACGACATAGCATTTTCTCCGTACGGTGAATACTGGAGA 414
QY 401 CAGATGAAGAAAATTTGGATCTTGAATTTCTGAGTGCACAAAAGTTCAATCTCCAGG 460
Db 415 CAAATTCGTAATAATTTGCAATTTGGAACCTTTGAGTGCACAAAGATGGTCAAAATTTTTTATAGC 474
QY 461 TTAATCCGAGGAGGAATAATGGAGATGCCATCAATTTCTCCGTTCCGAAAGCCGATCT 520
Db 475 TCGATTCCCAAGATGAGCTCTCGAAGATGCTCTCATCTATACGAACGACACCCCAATCTT 534
QY 521 CGGTCATATTTACAAAGATCATTTATGGCAATTAATAATTTGATCATGATAAGAAACATCC 580
Db 535 ACAGTCAATCTTACGACAAAATTTTGTGTTTACAGTTCCGTAACTTGTAGATCAGCT 594
QY 581 GTTGATAA---TTGTAAGCAAAAAGAAATGCTGAGTGTGCGAGTGCAGTCAATGAG 637
Db 595 TTAGGGAAGATATGTTGGTGAACGACAAATTTGATCATTTTTATGAGGGAATAATAATCA 654
QY 638 GCAGCGAGGATTTGGCACCGCAGACGCTTTTCGAGCTGGAATTAATTTACTTCACATATATC 697
Db 655 TTGGCAGGTGGAATTTAGTATGCTGATTTTTTCCCTACATGGAATAATGATTCATGATAT 714
QY 698 ATTTGAGCTGAGTCAAAACCCAGGCGTTTTCATCAGGAGATTCAGCATATCTTGAAGAG 757

Db 715 GATGGTTGGAATCTTAAACTGGTGAAGACACATCGTAGATTGATGAAATTTTGGGAAT 774
Qy |||||
Db 758 ATTCTTAATGAACACAAA-----GCCAATAGCCTTTTGAA 793
Db 775 GTTGTGATGAGCACAAAAGAACACAGACAGATGGCAAGAGGTAATGGTGAATTTGGT 834
Qy 794 GCGGATAACTTAATGGATGTTCTATTGAATCTTCAAAAAAATGGAAACGTTCCAGTGCCA 853
Db 835 GGTGAAGATTTGATGATGATTGTTAAGAGTTAGAGAAAGTGGAGAAGTTCAAATTCCT 894
Qy 854 GTGACAAACGAAGCATCAAGCATCCGTTTTCGAAATGTTTACCTGCGGGAGCGAAACA 913
Db 895 ATCAAAATGACAATATCAAAATCAATTAATTAATCGACATGTTCTCTGCGGGATCTGAACA 954
Qy 914 ACTTCGAAAGCTACAGAATGGGTAAATGGCAGAGCTGATGAAAAATCCAACTGAATCAAGA 973
Db 955 TCATCGACGACTATAATTTGGGCAATTAGCTGAATGATGAAGAAACCAAGTGTGTTAGCA 1014
Qy 974 AAAGCACAAGAGAAGTTAGACAAGTATTGGTGAATGGGAAAAAGTTGATGAATCAAGA 1033
Db 1015 AAGGCACAAGCTGAAGTAAGGCAAGCTTTGAAGGAGAAAAAGGTTTTCACACAGATTGAT 1074
Qy 1034 TTTTCATGATTTGAAATCTTCAAGTTAGTGGTTAAAGAAACTCTAAGATTACATCCTCG 1093
Db 1075 CTTGATGAGCTAAAAATATCTCAAGTTAGTAAATCAAGAAACCTTAAGAAATGCACCTCCA 1134
Qy 1094 GT---TGTCTTGATTCGAGGAGGTGTAGAGAAACAAACGAATTTGATGATATGAAATT 1150
Db 1135 ATCTCTTATTAGTTCCTAGAGAAATGATGGAGGATACAAGATTGATGGTTACAATATA 1194
Qy 1151 CATCCGAACACTCGAATTTGTAATGCTTGGGCGATAGGAAGAGATCCTAATACTTGG 1210
Db 1195 CCTTTCAAAACAAGAGTCATAGTTAATGCATGGCAATCGGACGAGATCCAGAAAGTTGG 1254
Qy 1211 TCGGAACCTGGAAAGTTTAAACCAGAAAGTTTAAAGATTGCGAATTTGATTAAGGG 1270
Db 1255 GATGACCCCGAAAGCTTTATGCCAGAGAGATTGAGAATAGTTCTATTGACTTCTTGGGA 1314
Qy 1271 ACGACATTTGAACCTGGTACCAATTTGGTCAGGAAAGAAATATGCTGGCATTACTTCA 1330
Db 1315 AATCATCATCAGTTTATACCATTTGGTCAGGAAAGAGATTGTCGGGAATGCTATTT 1374
Qy 1331 GCTATTACCAATTTGGAGTATGTCATTATATAAATCTATTATATCAATTTAAATGGGAAC 1390
Db 1375 GGTTTAGCTAATTTGGACAACCTTTAGCTCAGTTACTTTATCACTTCGATTGGAACTC 1434
Qy 1391 GCCGATGGAAATTACACCTCAAAACACTTGATATGACTGAAGCTATTGGGGGTGCTCTCAG 1450
Db 1435 CCTAATGGACAAAGTCATGAGAATTTCGACATGACTGACTCACCTGGAATTTCTGCTACA 1494
Qy 1451 AAAAAAATAGATCTTAAGTTGATTCCTATTCCATATCA 1488
Db 1495 AGAAAGGATGATCTTGTGTTGATGGCACTCCTTATGA 1532

Search completed: May 31, 2006, 11:50:40
Job time : 10159 sec8

THIS PAGE BLANK (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.
OM nucleic - nucleic search, using sw model
Run on: May 31, 2006, 09:01:43 ; Search time 8717 Seconds
(without alignments)
11117.151 Million cell updates/sec
Title: US-10-759-813-1
Perfect score: 1733
Sequence: 1 gcataaaaggaaatggagc.....tcttttcaatccgaaaaa 1733
Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 48236798 seqs, 27959665780 residues
Total number of hits satisfying chosen parameters: 96473596
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est3:*
3: gb_est4:*
4: gb_est5:*
5: gb_est6:*
6: gb_hnc:*
7: gb_est2:*
8: gb_est7:*
9: gb_est8:*
10: gb_est9:*
11: gb_gss1:*
12: gb_gss2:*
13: gb_gss3:*
14: gb_gss4:*
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match	Description
1	371.2	21.4	DV154738 CV03095B2
2	366.2	21.1	DV153215 CV03091A2
3	347.2	20.0	DV150019 CV03081B1
4	342	19.7	DV134278 CV03108A1
5	336.4	19.4	DV131796 CV03061A1
6	332.8	19.2	DV137266 CV03116B1
7	304.6	17.6	DV137008 CV03116A1
8	294.6	17.0	DN487910 R014G04.3
9	293	16.9	DV145886 CV03068A1
10	286.4	16.5	DV138221 CV03119A2
11	275.2	15.9	DF506382 WS0189.C2
12	274.2	15.8	DV119846 CV03024B1
13	273.2	15.8	902 8 CX053093
14	271.8	15.7	DV128282 CV03050B1
15	270	15.6	9 CX636845
16	270	15.6	9 CX670312
17	266.8	15.4	9 CX636844
18	261.6	15.1	794 10 DV124816
19	261	15.1	718 10 DV122991

20	258.8	14.9	742	10	DV148280
21	255.8	14.8	849	9	CX669008
22	255.8	14.8	900	9	CX663419
23	254.8	14.7	903	9	CX668839
24	253.8	14.6	731	10	DV126517
25	252.2	14.6	614	10	DV147676
26	252.2	14.6	872	10	DM003125
27	247.8	14.3	675	10	DV143418
28	247.8	14.3	799	10	DV127937
29	247.6	14.3	709	4	CA921009
30	245.2	14.1	742	10	DV153269
31	245.2	14.1	1016	9	CX701446
32	245	14.1	730	4	CA922231
33	243.6	14.1	738	10	DV136837
34	242.2	14.0	780	10	DV149902
35	242.2	14.0	834	10	DT730556
36	241.8	14.0	790	10	DV697990
37	240.8	13.9	902	8	CO125331
38	239.6	13.8	830	10	DR925887
39	237	13.7	692	10	DV144641
40	237	13.7	720	1	AI730111
41	236.8	13.7	599	7	BE204783
42	236.8	13.7	689	2	BM063831
43	236.8	13.7	994	9	CX666031
44	235.4	13.6	711	4	CA921435
45	235	13.6	918	10	DR933028

ALIGNMENTS

RESULT 1
DV154738
LOCUS
DEFINITION
CV03095B2F04.f1 CV03-normalized library Euphorbia esula cDNA clone
DV154738 654 bp mRNA linear EST 03-OCT-2005
DV154738.1 GI:76863745
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Euphorbia esula (leafy spurge)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Euphorbiaceae; Euphorbioideae;
Euphorbiaceae; Euphorbia.
REFERENCE
AUTHORS
Kim, W.R. and Mikel, M.
Direct submission, Anderson, J.V. 2005
TITLE
JOURNAL
COMMENT
Unpublished (2005)
Contact: James V. Anderson
USDA/ARS
1605 Albrecht Blvd, Biosciences Research Laboratory, Fargo, ND
58105, USA
Tel: 701 239 1263
Fax: 701 239 1252
Email: andersjv@fargo.ars.usda.gov
Adaptors and tags in 5'-end sequenced clones:
(Vector). . .TAAGCTTGATATCG(End Vector) (Start
EcoRI adaptor)AATTCATTGTGTGGG (End EcoRI adaptor) (Start
Insert). . .AAAAAAAAAAAAAAAAAAAA(End Insert) (Start Tag)TGGGT (End
Tag) (Start NotI site/Vector)GGCGGCCGCCGCGG. . . .Base
Calling/Quality Scores: PHRED from Washington University Genome
Center. Vector trimming: Cross_match from Washington University
Genome Center PHRAP suite. Low quality bases (Phred score < 20)
were trimmed from both ends of the sequence by an in-house script.
This sequence is vector free and at least 200 bp in length.
PCR Primers
FORWARD: TAATACGACTCACTATAGG(T7)
BACKWARD: ATTAAACCTCACTAAAG(T3)
Insert Length: 654 Std Error: 0.00
Plate: CV03095B2 row: F column: 04
Seq primer: TAATACGACTCACTATAGG (T7)

```
High quality sequence stop: 654.
Location/Qualifiers
1. 654
/organism="Euphorbia esula"
/mol_type="mRNA"
/db_xref="taxon:3993"
/clone="CV03095B2F04.f1"
/lab host="DH10B"
/clone lib="CV03-normalized library"
/insert="Vector: pBS II SK(+); Site 1: EcoRI(5' side of
tissue (leaf, stem, flowers, seeds, shoots, galls, root
and crown buds) from outdoor plants, cold acclimated
tissues (leaf, stem, and crown buds), crown and root buds
from decapitated greenhouse p ants, and seedlings. The
library was constructed and normalized as described by
Bonaldi, M.F., Lennon, G. and Soares, M.B. (1996), Genome
Research 6(9): 791-806. An identifying tag was added at the
3'during cDNA synthesis: insertAAAAAAAAAAAAAAAAATCGGT."

```

ORIGIN

```

Query Match      21.4%; Score 371.2; DB 10; Length 654;
Best Local Similarity 75.8%; Pred. No. 3.4e-84;
Matches 473; Conservative 0; Mismatches 148; Indels 3; Gaps 1;

QY 57 TTCTGCTGTTTAAATCTTAGTAGTATGATGAGGTGTGGAGAAACAGAAATCCACCTC 116
DB 31 TTCTAAATATCTTCACTTATTAAAGTTAAACATTAAGAACATGACTCAAAATGCTCCTC 90
QY 117 CAGGCCCATGGAAGTTCTATCATAGTAATCTTCCTCATTTTACTTACTCTCTGATC 176
DB 91 CAGGCCCATGGAAGTTTCAATATTAGGTAACTTCTCCTCATTTTACTTACTCTGATC 150
QY 177 TAGGCCATGAAGTTTATAGAGCTTGGCTCAAAATTTATGGACCTGTATGATCTTCAA 236
DB 151 AAATCATCAAGCTTTAGCTCTTATCTCTAAATTTATGGACCTATTATGATCTTCAAC 210
QY 237 TTGGCCAAAGTTTTCAGCTGTGTCATTTCTTCAGCTGAAGCAGCCAAAGAGTTTATGAAA 296
DB 211 TCGGCCAAGTCCAAATATTGTCATTTCTTCGGCTCAAGCAGTTAAAGAGTTTATGAAA 270
QY 297 CTCAGCTGATGCTTTCGCCCAACGCCCTATCGTCTTGGACCGCAGATGTGTTTTATA 356
DB 271 CTCATGTTGAATCTTCGCAGAACGTTCTCCGGGCTAGATCGCAGATTTGTTTTATG 330
QY 357 ATCGGAAGATGCTTTGTTGCTTCATATGGAGATCACTGGAGGCGAGATGAAGAAATTT 416
DB 331 ATAGGAATGATATCTTTTGTGTCATATGGAAATCATTTGGAGGCAATGAGAAATTT 390
QY 417 GGATACCTTGAATTTCTGATGTCGCAAAAAGTTTCAATCTCCAGGTTAATCCGAGAGGAAG 476
DB 391 GGATACCTGAATTTTATGTCGTAAGCAGATTTCAATCTTCAGGTTAATCCGGAGGAAG 450
QY 477 AAATGAGATGCGCATFACATTTCTCCGTTGGAAGCCCGGATCTCCGTCATATATACAA 536
DB 451 AAGTTTCGGAGGCGCATCAAAATCTCCTCGTTTCAAAAAGTGGATCTCCGTCATATCTTACAA 510
QY 537 AGATCAATTTATGGCATTAATTTTCGATCATGATAGACATCCGTTGCT---AATTGTA 593
DB 511 AAGTCATATTTGAGTTTCAAAATTTCAATCATGATAGAAACAATCATTTGTAAGAAATGTC 570
QY 594 AGCAAAAAGAAAGATTTGCTGAGTTGTCGGATGTCAGTCAATGAGGCGAGCAGTTTTTG 653
DB 571 ACCAAAAGAAAGACTTATGATGATTTGTTGGCGGTGTGAATGAGGCAATTTTCGAGTTTCG 630
QY 654 GCACCGCAGACGCTTTTCCGAGCT 677
DB 631 GGATTCGGATGCTTTTCCGTCGT 654
```

RESULT 2
DV153215

LOCUS DV153215 732 bp mRNA linear EST 03-OCT-2005
DEFINITION CV03091A2E05.f1 CV03-normalized library Euphorbia esula cDNA clone
ACCESSION CV03091A2E05.f1 5, mRNA sequence.
VERSION DV153215
KEYWORDS DV153215.1 GI:76862222
SOURCE EST.
ORGANISM Euphorbia esula (leafy spurge)
Euphorbia esula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Euphorbiaceae; Euphorbioideae;
Euphorbiaceae; Euphorbia.
1 (Bases 1 to 732)
Anderson, J.V., Horvath, D.P., Thimmapuram, J., Liu, L., Hernandez, A.,
Kim, W.R. and Mikel, M.
Direct submission, Anderson, J.V. 2005
Unpublished (2005)
Contact: James V. Anderson
USDA/ARS
1605 Albrecht Blvd, Biosciences Research Laboratory, Fargo, ND
58105, USA
Tel: 701 239 1263
Fax: 701 239 1252
Email: andersjv@fargo.ars.usda.gov
Adaptors and tags in 5'-end sequenced clones:
(Vector) . . . TAAGCTTGATCG (End Vector) (Start
EcoRI adaptor) AATTCATTGTGTGG (End EcoRI adaptor) (Start
Insert) . . . AAAAAAAAAAAAAAAAAA (End Insert) (Start Tag) TCGGT (End
Tag) (Start NotI site/Vector) GCGCGCCACCGCGG . . . Base
Calling/Quality Scores: PHRED from Washington University Genome
Center. Vector Trimming: Cross_match from Washington University
Genome Center PHRAP suite. Low quality bases (Phred score < 20)
were trimmed from both ends of the sequence by an in-house script.
This sequence is vector free and at least 200 bp in length.
PCR Primers
FORWARD: TAATCAGCTCACTATAGG(T7)
BACKWARD: ATTAACCTCACTAAG(T3)
Insert Length: 732 Std Error: 0.00
Plate: CV03091A2 row: E column: 05
Seq primer: TAATCAGCTCACTATAGG (T7)
High quality sequence stop: 732.
Location/Qualifiers
1. 732
/organism="Euphorbia esula"
/mol_type="mRNA"
/db_xref="taxon:3993"
/clone="CV03091A2E05.f1"
/lab host="DH10B"
/clone lib="CV03-normalized library"
/note="Vector: pBS II SK(+); Site 1: EcoRI(5' side of
insert); Site 2: NotI (3' side of insert); Tissue:
Seasonal crown buds from outdoor plants, whole plant
tissue (leaf, stem, flowers, seeds, shoots, galls, root
and crown buds) from outdoor plants, cold acclimated
tissues (leaf, stem, and crown buds), crown and root buds
from decapitated greenhouse p ants, and seedlings. The
library was constructed and normalized as described by
Bonaldi, M.F., Lennon, G. and Soares, M.B. (1996), Genome
Research 6(9): 791-806. An identifying tag was added at the
3'during cDNA synthesis: insertAAAAAAAAAAAAAAAAATCGGT."

FEATURES
source

ORIGIN

Query Match 21.1%; Score 366.2; DB 10; Length 732;
Best Local Similarity 70.3%; Pred. No. 6.8e-83;
Matches 507; Conservative 0; Mismatches 208; Indels 6; Gaps 1;

QY 587 AATTCTAGCAAAAGAAAGATTGCTGAGTGTTCGGATGCAGTCAATGAGCAGCGAGC 646
DB 12 AAATGCAAGAATCAAGAGCGGTGATGATGTGATTGATCGAGTGTGAGCTCGAGCT 71
QY 647 AGTTTGGCAGCGCAGACGCTTTTCCGACGTGGAATTTACTTCACTATATCAATTGAGCT 706
DB 72 GGGTTGAGTGTGCGGACGATTTTCCGTCGCTAAATTTCTTCACTATCTGAGTGAGAA 131

```

QY 707 GAGTCAAAACCCAGGCGTTTGCAATCAGGAGATTGACGATATACCTTGAAGAGATTCTTAAT 766
D 132 AAGTCAAAAGTTGCAGAAAGTTGCATAAAGACAGATGAGATTCTTGAAGAGATTATAAGT 191
QY 767 GAACAAAGCCCAATAGCCTTTTG-----AAGCGATAACTTAATGGATGTTCTATTG 820
D 192 GAACATAAGACTAATGCTAAGATTGGAAAGTCAAGCTGATTAATCTTTTGGATGTTTGTG 251
QY 821 AATCTTCAAAAAAATGGAACGTTCCAGTGCAGTGACAAACGAAAGCATCAAAAGCATCC 880
D 252 TAATCTCAGAAAATGGAATCTTCAAGTTCCATTGACTAATGATATATCAAGAGCT 311
QY 881 GTTTTGCAAATGTTACTGCGGAGCGAAACAACTTCGAAAGCTACAGAAATGGGTAATG 940
D 312 ACTCTGGAATGTTTCGGAGCTGGAGCGACACATCTCTCAAAAACCTACAGACTGGCAATG 371
QY 941 GCAGAGCTGATGAAAATCAACTGAATCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1000
D 372 GCGCAACTAATGAGGAAGCCATCAACATGAAAAGGCCCAAGAGAGGTTAGCGGGTC 431
QY 1001 TTTGGTGAATGGGAAAGTTGATGAATCAAGATTTCATGATTTGAAATCTTCAAGTTA 1060
D 432 TTTGGCGATACGGGAAGGTAGAAATCAAGAAATCAAGAACTCAAGAACTGAAATACTTGAATTA 491
QY 1061 GTGGTTAAAGAACTCTAAGATTACATCTCTCCGGTTGCTTGTGATTCGAGGGAGTGTAGA 1120
D 492 GTTGTAAAGAAACATTTGAGATTACATCTCTCCGGTGGCATTTGCTTGAAGATGCCGA 551
QY 1121 GAAACAAACGAATGATGATGAAATTCATCCGAACTCGAATGTTGTGAATGCT 1180
D 552 GAGAAAACCTAAAATTTGACGATATGATGTTTATCTTAAACCAAAAATCTTGTGAATCTCT 611
QY 1181 TGGCGGATGAGAGAGATCTTAATCTTGTGTCGGAACCTCGAAAGTTTACCCAGAAAGG 1240
D 612 TGGCGGATTTGGAAGATCCCAAGTTTGACCGACCCGGAAAGTTTCAACCCAGAAAGA 671
QY 1241 TTTAAAGATTGTGCAATGATTAATAAGGACGACATTTGAACTGGTACCATTTTGGTGA 1300
D 672 TTTCAAGATAGCCCAATAGACTACAAAGGTACAAATTTGCAATTTGATTCGTTGGTGTCT 731
QY 1301 G 1301
D 732 G 732

RESULT 3
DV150019
LOCUS
DEFINITION
CV03081BIH01.fl CV03-normalized library Euphorbia esula cDNA clone
CV03081BIH01.fl 5, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Euphorbia esula (leafy spurge)
Euphorbia esula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Euphorbiaceae; Euphorbiales;
Euphorbiaceae; Euphorbia.
1 (bases 1 to 696)
REFERENCE
AUTHORS
Anderson, J.V., Horvath, D.P., Thimmapuram, J., Liu, L., Hernandez, A.,
Kim, W.R. and Mikel, M.
Direct submission, Anderson, J.V. 2005
Unpublished (2005)
CONTACT: James V. Anderson
USDA/ARS
1605 Albrecht Blvd, Biosciences Research Laboratory, Fargo, ND
58105, USA
Tel: 701 239 1263
Fax: 701 239 1252
Email: andersjv@fargo.ars.usda.gov
Adaptors and tags in 5'-end sequenced clones:

```

```

(Vector) . . . TAAGCTTGATATCG (End Vector) (Start
EcoRI adaptor) AATTCATTTGTTGGG (End EcoRI adaptor) (Start
Insert) . . . AAAAAAAAAAAAAAAAAA (End Insert) (Start Tag) TGCCT (End
Tag) (Start NotI site) (Vector) GCGCGCCGACCGCGG. Base
Calling/Quality Scores: PHRED from Washington University Genome
Center. Vector trimming: Cross match from Washington University
Genome Center PHRAP suite. Low quality bases (Phred score < 20)
were trimmed from both ends of the sequence by an in-house script.
This sequence is vector free and at least 200 bp in length.
PCR Primers
FORWARD: TAATACGACTCACTATAGG (T7)
BACKWARD: ATTAACCCCTCACTAAAG (T3)
Insert Length: 696 Std Error: 0.00
Plate: CV03081BI row: H column: 01
Seq primer: TAATACGACTCACTATAGG (T7)
High quality sequence stop: 696.
FEATURES
Location/Qualifiers
1. .696
/organism="Euphorbia esula"
/mol_type="mRNA"
/db_xref="taxon:3993"
/clone="CV03081BIH01.fl"
/lab_host="DH10B"
/notes="Vector: pBS II SK(+); Site 1: EcoRI (5' side of
insert); Site 2: NotI (3' side of insert); Tissue:
Seasonal crown buds from outdoor plants, whole plant
tissue (leaf, stem, flowers, seeds, shoots, galls, root
and crown buds) from outdoor plants, cold acclimated
tissues (leaf, stem, and crown buds), crown and root buds
from decapitated greenhouse plants, and seedlings. The
library was constructed and normalized as described by
Bonaldi, M.F., Lennon, G. and Soares, M.B. (1996). Genome
Research 6 (9): 791-806. An identifying tag was added at the
3' during cDNA synthesis: InsertAAAAAAAAAAAAAAAAATCGCT."

```

ORIGIN

```

Query Match      20.0%; Score 347.2; DB 10; Length 696;
Best Local Similarity 69.0%; Pred. No. 5.3e-78;
Matches 475; Conservative 0; Mismatches 213; Indels 0; Gaps 0;

QY 831 AAATGGAACGTTCCAGTGCAGTCACAAACGAAAGCATCAAAAGCATCCGTTTTGCAAA 890
D 1 AACATGGAGATCTTCAACTTCGTTGACTGATGAAGTCATCAAAAGCCAATGTTATGAAA 60
QY 891 TGTTTACTGCCGGAGCGAAACAACTTCGAAAGCTACAGATGGTAAATGGCAGAGCTGA 950
D 61 TATTTTTCGGCGGAGTCATACATCTTCCATAACTGTGGAATGGGCTATGTCGGAGTTGA 120
QY 951 TGAATAATCCAACCTAAAGAAAGAAAGCAAGAAAGTTAGACAAGTATTTGGTGAAA 1010
D 121 TGAGGAGCCCAACATCAATGAGNAAGACACAGAGAGGTGAGACGAGTGTGGCAAGA 180
QY 1011 TGGGAAAAGTTGATGAATCAAGATTTTCATGTTGAAATTTCTTCAAGTAGTGTGTTAAAG 1070
D 181 AAGGAAAAGTCGAGGAGTCAAGAAATTCAGAAATGGAATACCTCATGTCAATGTCAAAG 240
QY 1071 AAATCTAAGATTACATCTCTCGGTTGCTTGTGATTCGGAGGAGTGTAGAAACACAC 1130
D 241 AAATCTTGAAGTTACACCTCCATCTACATTTGTTACTAGAGAAATCAGAGTAAAAACAA 300
QY 1131 GAATTCATCGATATGAAATTCATCCGACACATCGAATTTGTTGAAATGCTTGGCGCATAG 1190
D 301 AGTCAATGATATGATATTCACCCAGACTGTACTTCAATGTTATGATATGGCAATCG 360
QY 1191 GAAGAGATCCTAATACTTGTGTCGGAACCTCGGAAGTTTAAACCCAGAAAGGTTTAAAGATT 1250
D 361 GAAGAGACCCCGATGTTTGGATTGATGCGGAAAAGTTTATCCAGAAAGATTTGAAGATG 420
QY 1251 GTGCAATTCATATAAAGGAGCAGATTTGAACTGGTACCATTTGTCGAGGAAAAGAA 1310
D 421 GTCAATAGATTACAAAGGTGCAAAATATGGAATTAATACCGTTTGGTTCAGGAAAAGAA 480

```



```
Db 148 ATGGCGGTGACTCCCAATTGGTAAGAATATGGGATCAAGAAGCTTTGTTAAGTATAAAT 207
QY 623 GATGCAGTCAATGAGCGACGAGCTTTTGGCACCGCAGACGCTTTTCGCGTGGAA 682
Db 208 GATGGTGGAGTTGAGGGAGTGGCGGTATAGCATTTGCTGACGTGTTCTTCCTTTAA 267
QY 683 TTACTTCACTATACATTTGAGCTGAGTCAAAACCCAGGCGTTTGTGATCAGGAGATTGAC 742
Db 268 TTCTCTTCAATACATCACCGGAGAAATCTCGAAACTGATCCGGTTGCAATGAAGCGGAC 327
QY 743 GATATACCTTGNAGAGATCTTAAATGAACAAAGCCAAATAG----- 784
Db 328 AAGATCTTGAAGAGATCTTAAACGAAAGTAAAGAGAGAGAGAGAGAGTGGAGATCGTCGT 387
QY 785 CTTTTTGAAGCGGATAAATTAATGAGTGTCTATTGAATCTTCAAAAAATGGAACGTT 844
Db 388 GGTGGCGAAATCGAAATCTTTTGGATGTTCTTTTGGATTTCCAGAAATGGAATCTT 447
QY 845 CCAGTCCAGTGACAAACGAAAGCATCAAGCATCCGTTTTCGAAATGTTTACTGCGCGG 904
Db 448 CCAGTCCCGTTGACGAATGCTAGCATCAAGGAGCCATTTCTGGAATGTTTGGAGCGGG 507
QY 905 AGCGAAACAACTTCGAAAGCTACAGAAATGGGTAATGGCAGAGCTGATGAAAAATCAACT 964
Db 508 AGTGACACATCTCTCAAGACTACAGAAATGGGCAATGGCGGAATTAATGAGGCATCCAATA 567
QY 965 GAACTAAGAAAGACCAAGAAAGTTAGACAAGTATTTGGTCAATGGGAAAGTTGAT 1024
Db 568 GAATGAGAAAGTGCAGAGAGGTGAGCGAGCTTTTGGTGAACCGGAAACGTTGAG 627
QY 1025 GAATCAAGATTTTCATGATTTGAAATTTCTTCAAGTTAGTGGTTTAAAGAAATCTTAAGATTA 1084
Db 628 GAATCAAGGCTTGATGAATCTGAACTGACTTGTGAAGTTGTTGAAAGAAACGCTGAGACTA 687
QY 1085 CATCTCCGGTGTCTTGATTTCCGAGGAGTGTAGAGAAACAAACAGCAATGATGATAT 1144
Db 688 CATCTCCGGGT-CATTGATTTCTAGAGNATCAGTGTGAGAGAGAGATCGGTGGATAT 746
QY 1145 GAAATTCATCCGAACACTCGAATTTGTTGAAATG 1178
Db 747 GATATTCATCCCAAGACTAAAGTATCGTGAACG 780
```

```
RESULT 8
LOCUS DN487910/c
DEFINITION R014G04.3pr Populus root cDNA library Populus tremula x Populus tremuloides cDNA clone R014G04 3', mRNA sequence.
ACCESSION DN487910
VERSION DN487910
KEYWORDS EST.
SOURCE Populus tremula x Populus tremuloides
ORGANISM Populus tremula x Populus tremuloides
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE 1 (bases 1 to 793)
AUTHORS Sterky,F., Bhalerao,R.R., Unneberg,P., Segerman,B., Nilsson,P., Bruner,A.M., Charbonnel-Campaa,L., Lindvall,J.J., Tandré,K., Strauss,S.H., Sundberg,B., Gustafsson,P., Uhlen,M., Bhalerao,R.P., Nilsson,O., Sandberg,G., Karlsson,J., Lundberg,J. and Jansson,S.
A Populus EST resource for plant functional genomics
Proc. Natl. Acad. Sci. U.S.A. 101 (38), 13951-13956 (2004)
15353603
Other_ESTs: R014G04, R014G04.5pr
CONTACT: Bo Segerman
UNION: Umea Plant Science Center, Department of Plant Physiology
UNION: Umea University
UNION: 901 87 Umea, Sweden
Tel.: +46 90 786 5279
Fax: +46 90 786 6676
Email: bo.segerman@plantphys.umu.se.
Location/Qualifiers
```

FEATURES

```
source
1..793
/organism="Populus tremula x Populus tremuloides"
/mol_type="mRNA"
/db_xref="taxon:47664"
/clone="R014G04"
/tissue_type="root"
/clone_lib="Populus root cDNA library"

ORIGIN
Query Match 17.0%; Score 294.6; DB 9; Length 793;
Best Local Similarity 65.6%; Pred. No. 28-64;
Matches 460; Conservative 0; Mismatches 239; Indels 2; Gaps 2;

QY 790 TGAAGCGGATTAATTAATGATGTTCTATTGTAATCTTCA-AAAAATGGAAGCTTCCAG 848
Db 775 TGAAGCGGAGGATCTTCTGGATGTTCTTTTGAATCTCCAGAAAAACACGGGACCTTGGAT 716
QY 849 TGCCAGTGACAAACGAAAGCATCAAGCATCCGTTTTCGAAATGTTTACTGCGCGGAGCG 908
Db 715 TCCCTTTGACACACAGACAGCATCAAGCAACAATCCTGGAACTATTTCGCGCTGGAGTG 656
QY 909 AAACAACCTTCGAAAGCTACAGAAATGGGTAATGGCAGAGCTGATGAAAAATCCAACTGAAC 968
Db 655 ATACATCATCAACATTAATGGAATGGAACAATGGCCGAAATGTTAGAAATCCGATCGTAA 596
QY 969 TAAGAAAGGACAAAGAAAGTTAGACAAGTATTTGGTCAATGGGAAAGTTGATGAAT 1028
Db 595 TCGAAAGGACAAAGAGGAGTGGAGCAAGTCTTTAGTAATACTAGAAATGTTGATGAAA 536
QY 1029 CAAGATTTTCATGATTTGAAATCTTCAAGTTAGTGGTTAAAGAAATCTTAAGATTAATC 1088
Db 535 CATGCTTTCACAATTTGGAATTTCTTGAAGCTGATTAATCAAGAAATCTGAGACTGCATC 476
QY 1089 CTCGCGTGTCTTGATTTCCGAGGAGTGTAGAGAAACAACAAGAAATGATGATATGAAA 1148
Db 475 CTCAGCTCCTTTGATTTCAAGAGAAATGCAACAAGACGTCGAGATTAATGATATGTCA 416
QY 1149 TTCAATCCGAACACTCGAATTTGTTGAAATGCTTTGGCGGATAGGAAGAGATCTTAATCT 1208
Db 415 TACAGGCTCAAAAGTAAAGTATGATCAATCGTGGGCTATTGGAAGAGATTTCTGATCAT 356
QY 1209 GGTCCGAACTCGGAAAGTTTAAACCAGAAAGGTTTAAAGATTTGTGCAATTTGATTAAG 1268
Db 355 GGACTGAAAGCTGAGAAATTTCTATCCAGAGAGATTTCTAGACAGTTTCGATTGATTAAG 296
QY 1269 GAGCAGCATTTGAACTGGTACCATTTGTTGCGAGGAAAGAAATATGTCCTGCGATTAAT 1328
Db 295 GTACTTAATTTGAAATTCATCCCATTTGGTGTCTGGGAGAGGATGTCTCTGCAATTTAT 236
QY 1329 CAGCTATTACCAATTTGGAGTATGTCATTATAAATCTATTATATCATTTTAATTCGGAAC 1388
Db 235 TTGGTATAGCTACTGTCGAGCTTCACATTGGCAGCTGCTATACCATTTTCGATTGAAAC 176
QY 1389 TGGCCGATGGAATTAACACTCAAAACATTTGATATGACTGAAGCTATTGGCGGTGCTCA 1448
Db 175 TTCCAAACGAGAGACCTTTTGGAGATCTCGACATGAATGAAATTTTTCGGGCGACGGTTA 116
QY 1449 GAAAAAATATAGATCTTA-AGTTGATTTCTATTTCATATCA 1488
Db 115 GAAGGAAACATCAACTTACAGCTGATTCCTCATTTCCCTTTT 75

RESULT 9
LOCUS DVI45886
DEFINITION CV03068A1A02.f1 CV03-normalized library Euphorbia esula cDNA clone
ACCESSION DVI45886
VERSION DVI45886.1
KEYWORDS EST.
SOURCE Euphorbia esula (leafy spurge)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```



```

QY 885 TGCAAATGTTTCTGCGGGAGCGAAACAACCTTCGAAAGCTACAGAAGTGGTAAATGGCAG 944
Db 884 TGGAGATGTTTCTGCGGGAGTGAACCTCATCAACGATTTTGGAAATGGTACTGGCGG 765
QY 945 AGCTGATGAAATTCCTCACTGACTTAAGAAAGGACAGAGAAGTGTAGACAAGTATTG 1004
Db 764 AATTGATGAAATCCCAAGTGTAAATGGCAAGGACAGAAAGAGTGTAGGCAAGTATTGG 705
QY 1005 GTGAAATGGGAAAGTGTGATGAATCAAGATTTTCATGATTTGAAATTTCTTCAAGTTAGTGG 1064
Db 704 GTACTATTGGAGATGTTGATGAATCAAGTATCGTAAATTTGAAATTTCTTGAAGCTAATA 645
QY 1065 TTAAGAAACTCTAAGATTACATCTCCGGTTGCTTGAATCCGAGGAGTGTAGAGAAA 1124
Db 644 TCAAGAAACTTTTGAGATTTTCATCTCCAGGACCTTTTCATCCAGAGATGTGCGTCAATA 585
QY 1125 CAACAGAAATGATGGATATGAAATTCATCCGAACACTCGAATTG---TTGTGAATGCTT 1181
Db 584 GTTGTGAAGTTTCAACGATATGCCATAGAAGCTAACAGCAAAATTAGTAATGGTTAGTGCAT 525
QY 1182 GGGCGATAGGAGAGATCCTAATCTCTGTCGGAACCTTGGAAAGTTTAAACCCAGAAAGGT 1241
Db 524 GGGCTATTGGAAAGGATCCTAATTAATGATCGAAACCTGAGAAATTCCTACTGAGAGAT 465
QY 1242 TTAAGATTTGTGCAA TTGATTATAAAGGAGCAGACATTTGAACTGGTACCATTTGGTGCAG 1301
Db 464 TCCTAGATAGCTCGGTTGATTATAAGGTGCCAATTTTGAATTCATCCCATTTGGTCTG 405
QY 1302 GAAAGAAATATGCTCTGCGCATCTTCACTGATTAATACCAATTTGAGATGTGCTATTATA 1361
Db 404 GAAGAGAAATCTGCCCAGGAATGAACCTTTGGTTGGATAATGTGAGGTGGCTCTTGATA 285
QY 1362 ATCTATTATATCATTTTAAATGGGAACCTGGCCGATGGGAATACACCTCAAAACACTTGATA 1421
Db 344 RACTGTTATACATTTTGAATGGAACTTCCAATGGAGTAACACAGAGCTCTTGATA 285
QY 1422 TGACTGAAGCTATTGCGCGTGTCTTCAGGAAATAATAGATTTAAGTTGATTCCTATTTC 1481
Db 284 TGGAGGAGCATTTATAGTTCCCAACTAGAAAGACATGATCTTATATTAATTTCCCATCC 225
QY 1482 CATATC 1487
Db 224 CATATC 219

```

```

RESULT 12
LOCUS DV119846
DEFINITION DV119846 706 bp mRNA linear EST 30-SEP-2005
CV03024B1C05.f2 CV03-normalized library Euphorbia esula cDNA clone
CV03024B1C05.f2 5, mRNA sequence.
ACCESSION DV119846
VERSION DV119846.1 GI:76604495
KEYWORDS EST.
SOURCE Euphorbia esula (leafy spurge)
ORGANISM Euphorbia esula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Euphorbiaceae; Euphorbioideae;
Euphorbieae; Euphorbia.
1 (bases 1 to 706)
Anderson, J.V., Horvath, D.P., Thimmapuram, J., Liu, L., Hernandez, A.,
Kim, W.R., and Mikel, M.
Direct submission, Anderson, J.V. 2005
Unpublished (2005)
Contact: James V. Anderson
USDA/ARS
1605 Albrecht Blvd, Biosciences Research Laboratory, Fargo, ND
58105, USA
Tel: 701 239 1263
Fax: 701 239 1252
Email: andersjv@fargo.ars.ueda.gov
Adaptors and tags in 5'-end sequenced clones:

```

```

(Vector) . . . TAAGCTTGATATCG(End Vector)(Start
EcoRI adaptor)AATTCATGTTGGG(End EcoRI adaptor)(Start
Insert) . . . AAAAAAAAAAAAAA(End Insert)(Start Tag)TGCGT(End
Tag)(Start NotI site/Vector)GGCGCCACCGCGG . . . Base
Calling/Quality Scores: PHRED from Washington University Genome
Center. Vector Trimming: Cross match from Washington University
Genome Center PHRAP suite. Low quality bases (Phred score < 20)
were trimmed from both ends of the sequence by an in-house script.
This sequence is vector free and at least 200 bp in length.
PCR Primers
FORWARD: TAATACGACTCACTATAGGG(T7)
BACKWARD: ATTAACCCCTCACTAAAG(T3)
Insert Length: 706 Std Error: 0.00
Plate: CV03024B1 row: C column: 05
Seq primer: TAATACGACTCACTATAGGG (T7)
High quality sequence stop: 706.
FEATURES
Location/Qualifiers
1..706
/organism="Euphorbia esula"
/mol_type="mRNA"
/db_xref="taxon:3993"
/clone="CV03024B1C05.f2"
/lab_host="DH10B"
/clone_lib="CV03-normalized library"
/notes="Vector: pBS II SK(+); Site 1: EcoRI(5' side of
insert); Site 2: NotI(3' side of insert); Tissue:
Seasonal crown buds from outdoor plants, whole plant
tissue (leaf, stem, flowers, seeds, shoots, galls, root
and crown buds) from outdoor plants, cold acclimated
tissues (leaf, stem, and crown buds), crown and root buds
from decapitated greenhouse p ants, and seedlings. The
library was constructed and normalized as described by
Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome
Research 6(9): 791-806. An identifying tag was added at the
3'during cDNA synthesis: insertAAAAAAAAAAAAAAAAATCGT."

```

ORIGIN

```

Query Match 15.8%; Score 274.2; DB 10; Length 706;
Best Local Similarity 68.6%; Pred. No. 3.5e-59; Indels 0; Gaps 0;
Matches 375; Conservative 0; Mismatches 172;
QY 942 CAGAGCTGATGAAAAATCCAACTGAACCTAAGAAAGACAAAGAAAGTGTAGACAAGTAT 1001
Db 1 CAGAACTNNTNTACACCCAGAACCAATGAAAAAGCACAGATGAAGTACGACAAGTTT 60
QY 1002 TTGTGAAATCGGAAAAAGTTGATGAATCAAGATTTTCATGATTTGAAATTTCTCAAGTAG 1061
Db 61 TTGCAGAAAAAGGAAATGTAGAAGAAATCACAACTTCACCAATTTAAATTACTTTGAAATTAG 120
QY 1062 TGGTTAAGAACTCTAAGATTACATCTCCGGTTGTCTTGATTCGAGGGAGTGTAGAG 1121
Db 121 TCATCAAGAAACCCCTCAGATTACACCAGAGTCCCTTTAATCCAGAGAAATCCAGAG 180
QY 1122 AAACAACACGAATTTGATGGATATGAAATTCATCCGAACACTCGAATTTGTTGAATGCTT 1181
Db 191 AAAAAACAAAGGTCAATGATATGACATTTCTCCCAAAAACTAAGGCACCTTTGTAATGTTT 240
QY 1182 GGGCGATAGGAAGAGATCCTAATACTTGGTCGGAACCTTGGAAAGTTTAAACCCAGAAAGGT 1241
Db 241 GGGCAATTTCTAGGGACCCAAAAATTTGGCCCTGAAGACAGATAAAATTTATACCTGAAAGAT 300
QY 1242 TTAAGATTGTGCAATTTGATTATAAGGACGACATTTGMACTGTCACCATTTGGTGCAG 1301
Db 301 TTGAAAAATAGTTCATATAGATTTTAAGGGAATAATTTGGAATTCGCTCGCTTTGGTTGAG 360
QY 1302 GAAAAAGAAATATGCTCTGGCATTTACTTCAGCTATTACCAATTTGGAGTATGTCAATTATAA 1361
Db 361 GAAAAAGAAATATGTCAGGCATGGCGTTGGGGGTAACTAAATTTGGAGCTTTTCTTTCGCAC 420
QY 1362 ATCTATTATATCATTTTAAATTTGGGAACCTGGCCGATGGAATTAACACTCAAAACACTTGATA 1421
Db 421 AGCTTCTGTATCATTTTGACTGGAACTTGCCGATGGAAAGACATGTGTCGGGAACCTTGACA 480

```


ORGANISM Euphorbia esula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Malpighiales; Euphorbiaceae; Euphorbioideae; Euphorbieae; Euphorbia.

REFERENCE
1 (bases 1 to 674)
Anderson, J.V., Horvath, D.P., Thimmapuram, J., Liu, L., Hernandez, A., Kim, W.R. and Mikel, M.
Direct submission, Anderson, J.V. 2005
Unpublished (2005)
Contact: James V. Anderson
USDA/ARS
1605 Albrecht Blvd, Biosciences Research Laboratory, Fargo, ND 58105, USA
Tel: 701 239 1263
Fax: 701 239 1252
Email: andersjv@fargo.ars.usda.gov

Adaptors and tags in 5'-end sequenced clones:
ECORI adaptor: .TAAGCTTGATCG(End Vector) (Start
Insert) . . . AAAAAAAAAAAAAAAAAA(End Insert) (Start Tag) TCGGT (End Tag) (Start NotI site/Vector) GCGCGCCGACGCGG. . . Base
Calling/Quality Scores: PHRED from Washington University Genome Center. Vector Trimming: Cross_match from Washington University Genome Center PHRAP suite. Low quality bases (Phred score < 20) were trimmed from both ends of the sequence by an in-house script. This sequence is vector free and at least 200 bp in length.
PCR Primers
FORWARD: TAATACGACTCACTAGG(T7)
BACKWARD: ATTAACCTCTACTAAG(T3)
Insert Length: 674 Std Error: 0.00
Plate: CV03050B1 row: D column: 03
Seq primer: TAATACGACTCACTAGG (T7)
High quality sequence stop: 674.

FEATURES
Location/Qualifiers
1..674
/organism="Euphorbia esula"
/mol_type="mRNA"
/db_xref="taxon:3993"
/clone="CV03050B1003.f1"
/lab_host="DH10B"
/clone_lib="CV03-normalized library"
/note="Vector: pBS II SK(+); Site 1: EcoRI(5' side of insert); Site 2: NotI (3' side of insert); Tissue: Seasonal crown buds from outdoor plants, whole plant tissue (leaf, stem, flowers, seeds, shoots, galls, root and crown buds) from outdoor plants, cold acclimated tissues (leaf, stem, and crown buds), crown and root buds from decapitated greenhouse p ants, and seedlings. The library was constructed and normalized as described by Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome Research 6(9): 791-806. An identifying tag was added at the 3' during cDNA synthesis: insertAAAAAAAAAAAAAAAAAATGCGT."

source

RESULT 15

LOCUS CX636845

DEFINITION UCRPT02-UCR1 Poncirus trifoliata roots with Iron deficiency - mRNA sequence.

ACCESSION CX636845

VERSION CX636845.1 GI:57871674

KEYWORDS EST.

SOURCE Poncirus trifoliata

ORGANISM Poncirus trifoliata

REFERENCE 1 (Bases 1 to 819)

AUTHORS Close, T.J., Roose, M.L., Parker, D.R., Federici, C.F., Mandal, J., Fenton, R.D., Wanmaker, S., Landry, B., Hubert, N., Laforest, M., Landry, J. and Ligonde, A.

TITLE Development of EST Resources and New Genetic Markers for California Citrus - Poncirus trifoliata Roots with Iron Deficiency - UCRPT02-UCR1

JOURNAL Unpublished (2005)

COMMENT Contact: Timothy J. Close
Department of Botany & Plant Sciences
University of California
Riverside, CA 92521-0124, USA
Tel: 909-787-3318
Fax: 909-787-4437
Email: timothy.close@ucr.edu
Seq primer: T3.

FEATURES
Location/Qualifiers
1..819
/organism="Poncirus trifoliata"
/mol_type="mRNA"
/cultivar="Pomeyoy"
/db_xref="taxon:37690"
/clone="UCRPT02-22G04-M7-1-5.g"
/tissue_type="root"

Db 181 GCTATGTTGGAAAAAGTGAACAGCAGTCTTTAGAGACGAGTAGTGGTA---CGGATCTG 237
Qy 668 TTTCCGACGTGGAAATTAATCTTCACTATATCTGAGCTGATCAAAACCCAGCGCTTTG 727
Db 238 TTTCTCTCGTTGAAATTCATTTCAGTATATCAATGCTGAGAAAGTCAAGAATGCGCAAGGTA 297
Qy 728 CATCAGGAGATTGAGTATATCTTGAAGAGATTCTTAAATGAACACAAAGCCCAATAGCCT 787
Db 298 CAAAGGAAATGGATAAAATGCTTGAAGATATTATTAAGAAACACAAAGTTCAATATAAG 357
Qy 788 TTTGAAGCGGATAACTTAATGATGTTCTTATTTGAATCTTCAAAAAAATGGAACCTTCCA 847
Db 358 TTTGAGGATAACAATCTTTTACAGTTCTTTTGAATCAACAGCAAAATGGAGATCTTGA 417
Qy 848 GTGCCAGTGAACAAACGAAAGCATCCGTTTTCGAAATGTTTACTGCCCGGAGC 907
Db 418 CTTCATTCGACAAATGAAGTCATCAAGCCCAATATATGGAATATTTTTTGGTGGAGC 477
Qy 908 GAAACAATTCGAAAGCTACAGATGGTAATGCGAGCTGATCAAAAATCCAATCGAA 967
Db 478 CATACATCTTCTAAATTTGTGGAATGGGCAATGTCAGAGTTGATGAAGAACCCAGAAATCA 537
Qy 968 CTAAGAAAGCACAAAGAAAGTTAGACAAGTATTGCTGAAATGGGAAAGTTGATGAA 1027
Db 538 ATGACAAAGCACAAAGAGAGTGAACAATCTTTGTTGAAAGGGAAGTCCAGGAA 597
Qy 1028 TCAAGATTTCATGATTTGAAATTTCTCAAGTTAGTGGTTAAAGAACTCTAAGATTACAT 1087
Db 598 TCAATATGCAAGAACTGAAATACCTCAAGTCAATATCAAAAGAACTCTCAGATTGCAC 657
Qy 1088 CTTCCGG 1094
Db 658 CTTCCGG 664

ORIGIN
Query Match 15.7%; Score 271.8; DB 10; Length 674;
Best Local Similarity 64.0%; Pred. No. 1.4e-58;
Matches 427; Conservative 0; Mismatches 237; Indels 3; Gaps 1;
Qy 428 TTTCTGCTGCCAAAAGTTCAATCCCTCCAGTTAATCCGAGAGAGAAATGCGGAT 487
Db 1 TTTCTGAGTCTCGGTACAAAGCAATTCGTTAAATCTGACGAAATCCGTCGCT 60
Qy 488 GCATCACATTCCTCGTTTCGAAAGCCGGATCTCCGGTCAATATTACAAAGATCATTTAT 547
Db 61 TTTATCAGATTCCTCGGTACAAAGCAATTCGTTAAATCTGACGAAATCCGTCGCT 120
Qy 548 GCATTAATAATTCGATCATGATAGAACATCCGTTGGTAAATGTAAGCAAAAAGAAAGA 607
Db 121 GATTTGACAAAGTTCAATATGCTTTATTAACCTTATGGAATAATGTCATGAAAAATTTGTTG 180
Qy 608 TTGCTGAGTGTGCGCGATGCAGTCAATGAGGAGGACGAGTTTGGGACCCGACGCT 667

/dev_stage="seedling"
/lab_host="E. coli TJU121"
/clone_lib="Poncirus trifoliata Roots with Iron Deficiency - UCRT02-UCR1"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site 1: EcoRI; Site 2: XhoI; Plant materials were prepared by Federici (Roose lab) with advice from Parker. Seedlings 138 days after sowing were about 20-23 cm in height, bearing 8-12 leaves. On May 26, 2004, plants were washed free of soil with a stream of water, then placed upright with the roots submerged in two tanks of nutrient solution. The solution was sufficient in all major and minor nutrients and buffered with MES at about pH 6.95. The plants were maintained in this until June 14, 2004, when the solution was changed to one with only 20 micromolar iron, chelated with EDTA to induce iron deficiency. The pH was maintained at 6.99 by sodium carbonate/CO2 buffering. This solution was replaced on July 6, 2004. Roots from three plants were sampled on June 16, June 21, July 1 and July 28, 2004. Roots were collected by removing the plant from the nutrient solution, blotting off excess moisture with a paper towel, then cutting off the top of the plants. Three plants were pooled in one aluminum foil packet, and frozen between two sheets of dry ice. The time between removal from solution and freezing on dry ice did not exceed one minute. The frozen tissue was stored in these foil packs at -80°C. This sampling strategy did not correspond to initiation of stress in the plant, but only to when the plants entered the low iron solution. The actual stress was not initiated until the internal iron reserve was depleted. By the July 1 sampling date, slight iron deficiency symptoms were apparent. By the final sampling date, clear iron deficiency symptoms were present. By the July 1 sampling date, the roots showed growth of *Thielaviopsis brevicola*, a fungus. Other than what came away when the roots were blotted with paper towels, no effort was made to remove the fungus because it is not just a surface contaminant, but grows within the roots as well. Mandal and Fenton (Close lab) purified RNA using TRIzol, poly(A) mRNA using an Oligotex mRNA Kit (Qiagen), produced a primary cDNA library using a lambda ZAP XR cDNA Synthesis Kit (Stratagene), then mass-excised 0.5 million pfu from the primary library to produce a phagemid population. The library was made from equal portions of RNA from each of the four collection dates. Phagemids were plated, plasmid DNA purified, cDNA clones archived, and DNA sequences determined bi-directionally using an ABI3730 at DNA Landmarks (Landry, Hubert, Laforest, Landry, Ligonde). Chromatogram files were downloaded by FTP by Close, then processed by Mananaker (Close lab) using the HarVest pipeline (<http://harvest.ucr.edu>) to remove vector and cloning oligo sequences and various contaminants, and to trim to a high quality region. Sequences that retained a phred 17 region of at least 100 bases were assembled, then chimeras were removed following manual inspection of assemblies (Close, Roose, Mananaker). Sequences that survived all removal steps were submitted to GenBank."

ORIGIN

Query Match 15.6%; Score 270; DB 9; Length 819;
Best Local Similarity 60.7%; Pred. No. 4.4e-58;
Matches 460; Conservative 0; Mismatches 295; Indels 3; Gaps 1;
QY 791 GAAGCGGATACTTAATGATGTTCTATTGAATCTTCAAAAAATGGAAACGTTCCAGTG 850
DB 30 GAAGTTGAAGATCTTGTTCATGTTCTGTGAAAGCTCAAGAGGATGGTAATCTTGAATTT 89
QY 851 CCAGTGACAAACGAAGCATCAATCCGTTTTCGAATCTTTACTCGCGGAGCGNA 910
DB 90 CCCCTAACTACCAAAAACGTCAAAGCTGTTGTCATGGATGTTTTTAGTGCTGGCACCGAG 149
QY 911 ACAACTTCCAAAGCTACAGATGGGTAATGGCAGAGCTGATGAAAAATCCAACTGAACTA 970

Db 150 ACATCAGCAACAACCTGTAGATTGGGCTATGCTCTTGTGATGAAGTATCCAAAAATAATG 209
QY 971 AGAAAAACCAAGAAAGAGATTAGACAAGTATTTGGTGAATAATGGGAAAAAGTTGATGAATCA 1030
Db 210 AAAAAAGGCACAAGATGAGGTAAAGGAAGTCTTTTAATAGAAAGGGGAAAGTCGATGAAGCG 269
QY 1031 AGATTTTCATGATTTCGAAATTCCTCAAGTTAGTGGTTAAAGAAACTCTTAAGATTACATCCT 1090
Db 270 GGCCTTGAGGAATGAAGTACTTTGAAGCTAGTTTAAAGAGACTCTGAGAAATACAGCCT 329
QY 1091 CCGGTT---GTCTTGATTCGAGGGAGTGTAGAGAAAAACAACAGAAATTTGATGATGAATGAA 1147
Db 330 TCTGCTCCAATGTTATTTCCACAGAGATGTCAGAGACGTGTAATATTAAATGATACGAC 389
QY 1148 ATTCATCCGAACACTCGAATGTTGTAATGCTTGGGCGATAGGAAGAGATCCTAATACT 1207
Db 390 ATACAGCCCAAGCAAGTACTTGTTAATGTATGGCGGATGACAAGAGATTCCAAATAT 449
QY 1208 TGGTCGGAACTGGAAAGTTTAAACCAGAAAGGTTTAAAGATTGTGCAATTTGATTATAAA 1267
Db 450 TGGACTGAACCTGAGAGCTTTATGCGAGAGAGTTCCTTGATTCCTATAGACTTTAAG 509
QY 1268 GGGACGACATTTGAACTGCTACCATTTGGTGCAGGAAAAAGAAATATGCTCGCATTACT 1327
Db 510 GGTTCCTAATTTTGAATTTCTCCCATTTGGTCTGGTAGGAGGTGTGCCCTGGAAATGACA 569
QY 1328 TCAGCTNTTACCAATTTGGAGTATGTCATTATATAATCTATTATATCAATTTTAATTCGGAA 1387
Db 570 TATGCCAATGCCAACATTTGAGCTTCCACTTGCATTTGTTTGTACCAATTTCCGATTGGAAG 629
QY 1388 CTGGCCGATGGAATTCACACTCAACACTTTGATATGACTGAAAGCTATTGGCGGTGCTCTC 1447
Db 630 CTACCAATGGAAATGAAGATGAGGACTTAGACATGACAGAGGATTTGGTGTGACCGTC 689
QY 1448 AGGAAAAAATAGATCTTTAAGTTGATTCCTATTTCCATATCAAGTTAGCTTAGGCTCAAAAT 1507
Db 690 AAAAGAAAGCAAAATCTGTGCTGATTCCTTATATATCTCCAGCTGTGGCATAA 749
QY 1508 ATTTCTTGATTACATAGGAGGTTGAAATATATATAAT 1545
Db 750 GCTCTAGAAGGAAAAAACTACAGTATAATTTTGAATAAT 787

Search completed: May 31, 2006, 14:15:58
Job time : 8723 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 31, 2006, 08:41:59 ; Search time 1147 Seconds
(without alignments)
10534.356 Million cell updates/sec

Title: US-10-759-813-1

Perfect score: 1733

Sequence: 1 gcataaagggaatggagc.....tcttttcaaatccgaaaaa 1733

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_8.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004as.*
- 13: Geneseqn2004bs.*
- 14: Geneseqn2005s.*
- 15: Geneseqn2006s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1733	100.0	1733	6	Abk11137 CDNA enco
2	410	23.7	1538	12	Adn10568 Nicotiana
3	410	23.7	1538	14	Adz65256 Tobacco f
4	410	23.7	1538	14	Adz65803 Tobacco c
5	410	23.7	1538	15	Aee06775 Tobacco c
6	410	23.7	1538	15	Aef97168 Tobacco c
7	400.4	23.1	1673	12	Adn10646 Nicotiana
8	400.4	23.1	1673	12	Adn10624 Nicotiana
9	400.4	23.1	1673	12	Adn10626 Nicotiana
10	400.4	23.1	1673	14	Adz65334 Tobacco f
11	400.4	23.1	1673	14	Adz65312 Tobacco f
12	400.4	23.1	1673	14	Adz65314 Tobacco f
13	400.4	23.1	1673	14	Adz65861 Tobacco c
14	400.4	23.1	1673	14	Adz65859 Tobacco c
15	400.4	23.1	1673	14	Adz65881 Tobacco c
16	400.4	23.1	1673	14	Aee06853 Tobacco p
17	400.4	23.1	1673	14	Aee06831 Tobacco p
18	400.4	23.1	1673	14	Aee06833 Tobacco p

19	400.4	23.1	1673	15	Aef97224	Tobacco c
20	400.4	23.1	1673	15	Aef97226	Tobacco c
21	400.4	23.1	1673	15	Aef97246	Tobacco c
22	399.4	23.0	1781	4	Adn15839	Pepper cy
23	399.4	23.0	1781	7	Ado58604	Pepper cy
24	398.8	23.0	1673	12	Adn10628	Nicotiana
25	398.8	23.0	1673	14	Adz65316	Tobacco f
26	398.8	23.0	1673	14	Adz65863	Tobacco c
27	398.8	23.0	1673	14	Aee06835	Tobacco p
28	398.8	23.0	1673	15	Aef97228	Tobacco c
29	393.2	22.7	1610	12	Adn10676	Nicotiana
30	393.2	22.7	1610	12	Adn10612	Nicotiana
31	393.2	22.7	1610	14	Adz65300	Tobacco f
32	393.2	22.7	1610	14	Adz65364	Tobacco f
33	393.2	22.7	1610	14	Adz65911	Tobacco c
34	393.2	22.7	1610	14	Adz65847	Tobacco c
35	393.2	22.7	1610	14	Aee06819	Tobacco p
36	393.2	22.7	1610	14	Aee06883	Tobacco p
37	393.2	22.7	1610	15	Aef97276	Tobacco c
38	393.2	22.7	1610	15	Aef97212	Tobacco c
39	388.8	22.4	1566	12	Adn10606	Nicotiana
40	388.8	22.4	1566	14	Adz65294	Tobacco f
41	388.8	22.4	1566	14	Adz65841	Tobacco c
42	388.8	22.4	1566	14	Aee06813	Tobacco p
43	388.8	22.4	1566	15	Aef97206	Tobacco c
44	388.8	22.4	1576	12	Adn10564	Nicotiana
45	388.8	22.4	1576	14	Adz65252	Tobacco f

ALIGNMENTS

RESULT 1

ABK11137
ID ABK11137 standard; cDNA; 1733 BP.

XX ABK11137;

XX 05-JUN-2002 (first entry)

XX CDNA encoding Euphorbia lagascae cytochrome P450 enzyme.

XX Cytochrome P450 enzyme; synthesis of delta 12-epoxy fatty acid; epoxide;
KW vernolic acid; modified fatty acid; oil; commercial crop; plasticiser;
KW crosslinking coating application; setting printing ink; transgenic;
XX plant; EST; clone eallc_pk002_i4; gene; ss; expressed sequence tag.

XX Euphorbia lagascae.

XX Key Location/Qualifiers

FT CDS 14..1516
FT /tag= a
FT /product= "Cytochrome P450 enzyme"

XX WO200208269-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US022790.

XX 21-JUL-2000; 2000US-0219833P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Cahoon EB;

XX WPI; 2002-257331/30.

XX P-PSDB; AAU77470.

XX Novel isolated polynucleotide encoding plant cytochrome P450 enzyme
PT associated with synthesis of delta12-epoxy fatty acids, useful for
PT creating transgenic plants with higher or lower level expression of the
PT enzyme.

XX	Claim 8; Page 49; 53pp; English.	
PS	The present invention relates to the isolation of the polynucleotide and	
XX	polypeptide sequences for a plant cytochrome P450 enzyme associated with	
CC	the synthesis of delta 12-epoxy fatty acids from Euphorbia lagascae. The	
CC	polynucleotide sequence of the invention can be used to produce epoxide	
CC	containing fatty acids such as vernolic acid. The sequences of the	
CC	invention can be used to manipulate modified fatty acids to produce oils	
CC	in commercial crops. They can also be used to produce plasticisers, for	
CC	crosslinking coating applications, and setting printing inks. The	
CC	polynucleotide can also be used for creating transgenic plants in which	
CC	the enzyme is present at higher or lower levels than normal, in cell	
CC	types or in developmental stages in which they are not normally found.	
CC	The present sequence encoding the E. lagascae cytochrome P450 enzyme of	
CC	the invention is isolated from the expressed sequence tag (EST) clone	
CC	ee1ic.pk002.i4	
XX	Sequence 1733 BP; 556 A; 282 C; 376 G; 519 T; 0 U; 0 Other;	
SQL	Query Match 100.0%; Score 1733; DB 6; Length 1733;	
	Best Local Similarity 100.0%; Pred. No. 0;	
	Matches 1733; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 GCATAAAGGAAATGGAGCAGAAATCTCTCTTTCCGAGCATTTTAATAGTTTTCT 60	
DB	1 GCATAAAGGAAATGGAGCAGAAATCTCTCTTTCCGAGCATTTTAATAGTTTTCT 60	
QY	61 GCTTGTGTTTAACTTTAGTAGTAGTCATGAGGTTGTGGAAGAAACAGAAATCCACCTCCAGG 120	
DB	61 GCTTGTGTTTAACTTTAGTAGTAGTCATGAGGTTGTGGAAGAAACAGAAATCCACCTCCAGG 120	
QY	121 GCCATGGAGTTTCCCTCATAGTAGTAATCTTCTCATTTATTACTACCTTCTGATCTAGG 180	
DB	121 GCCATGGAGTTTCCCTCATAGTAGTAATCTTCTCATTTATTACTACCTTCTGATCTAGG 180	
QY	181 CCATGAACGTTTGTAGAGCCTTGGCTCAATTTATGACCTGTTATGAGTCTTCAAAATGG 240	
DB	181 CCATGAACGTTTGTAGAGCCTTGGCTCAATTTATGACCTGTTATGAGTCTTCAAAATGG 240	
QY	241 CCAAGTTTCAGCTGTTGTCATTTCTTCCAGTCAAGCAGCAGCAAGAGGTTATGAAAACTCA 300	
DB	241 CCAAGTTTCAGCTGTTGTCATTTCTTCCAGTCAAGCAGCAGCAAGAGGTTATGAAAACTCA 300	
QY	301 GGCTGATGCTTCGCCCAACGCCCTATCGTCTTGGACGACAGATTTGTTTTATAATCG 360	
DB	301 GGCTGATGCTTCGCCCAACGCCCTATCGTCTTGGACGACAGATTTGTTTTATAATCG 360	
QY	361 GAAAGATGCTTGTGTTGCTTCATATGGAGATCACTGGAGGCAGATGAAGAAATTTGGAT 420	
DB	361 GAAAGATGCTTGTGTTGCTTCATATGGAGATCACTGGAGGCAGATGAAGAAATTTGGAT 420	
QY	421 ACTTGAATTTCTGAGTGCCAAAGAAAGTTCAATCTCCAGGTTAATCCGAGAGAGAAAT 480	
DB	421 ACTTGAATTTCTGAGTGCCAAAGAAAGTTCAATCTCCAGGTTAATCCGAGAGAGAAAT 480	
QY	481 GGAGGATGCCATCACATTTCTCGGTTGAAAGCCGATCTCCGGTCAATATTACAAAGAT 540	
DB	481 GGAGGATGCCATCACATTTCTCGGTTGAAAGCCGATCTCCGGTCAATATTACAAAGAT 540	
QY	541 CATTTATGCGCATTAATTTTCGATCATGATAGTAAGACATCCGTTGGTGAATTTGAAGCAAAA 600	
DB	541 CATTTATGCGCATTAATTTTCGATCATGATAGTAAGACATCCGTTGGTGAATTTGAAGCAAAA 600	
QY	601 AGAAGATTTGCTGAGTTGTCGATGTCAGTCAATGAGGCGAGCGAGTTTGGCACCGC 660	
DB	601 AGAAGATTTGCTGAGTTGTCGATGTCAGTCAATGAGGCGAGCGAGTTTGGCACCGC 660	
QY	661 AGACGCTTTTCCGAGCGTGGAAATTAATCTCATATATCATTTGAGCTGAGTCAAAACCCAG 720	
DB	661 AGACGCTTTTCCGAGCGTGGAAATTAATCTCATATATCATTTGAGCTGAGTCAAAACCCAG 720	
QY	721 GCGTTTGCATCAGGAGATTGACGATATCTTGAAGAGATTCTTAATGAACACAAAGCCAA 780	
DB	721 GCGTTTGCATCAGGAGATTGACGATATCTTGAAGAGATTCTTAATGAACACAAAGCCAA 780	
QY	781 TAAGCCTTTTGAAGCGGATACTTAATGGAATCTTAAATCTTCAAAAAATGGAAA 840	
DB	781 TAAGCCTTTTGAAGCGGATACTTAATGGAATCTTAAATCTTCAAAAAATGGAAA 840	
QY	841 CGTTCCAGTCCAGTGACAAACGAAAGCATCAAGCATCCGTTTGGCAATGTTTACTGC 900	
DB	841 CGTTCCAGTCCAGTGACAAACGAAAGCATCAAGCATCCGTTTGGCAATGTTTACTGC 900	
QY	901 CGGAGCGAAACAACTTCCGAAAGCTACAGAAATGGTAAATGGCAGAGCTGATGAAAAATCC 960	
DB	901 CGGAGCGAAACAACTTCCGAAAGCTACAGAAATGGTAAATGGCAGAGCTGATGAAAAATCC 960	
QY	961 AACTGAACATAAGAAAAAGCAAGAAGAGTTAGACAAGTATTTGGTGAATGGGAAAAAGT 1020	
DB	961 AACTGAACATAAGAAAAAGCAAGAAGAGTTAGACAAGTATTTGGTGAATGGGAAAAAGT 1020	
QY	1021 TGATGAATCAAGATTTTCATGATTTTGAATTTCTTCAAGTTAGTGGTTAAAGAAACTCTAAG 1080	
DB	1021 TGATGAATCAAGATTTTCATGATTTTGAATTTCTTCAAGTTAGTGGTTAAAGAAACTCTAAG 1080	
QY	1081 ATTACATCTCTCCGGTTGCTTCTTCCGAGGAGTGTAGAGAAACAAACAGAAATGATGG 1140	
DB	1081 ATTACATCTCTCCGGTTGCTTCTTCCGAGGAGTGTAGAGAAACAAACAGAAATGATGG 1140	
QY	1141 ATATGAAATTCATCCGAACTCGAATTTGTTGTAATCTTGGCGATAGGAAGATGCC 1200	
DB	1141 ATATGAAATTCATCCGAACTCGAATTTGTTGTAATCTTGGCGATAGGAAGATGCC 1200	
QY	1201 TAATACCTTGGTCCGAACTCGAAAGTTTAAACCCAGAAAGGTTTAAAGATTGTGCAATTGA 1260	
DB	1201 TAATACCTTGGTCCGAACTCGAAAGTTTAAACCCAGAAAGGTTTAAAGATTGTGCAATTGA 1260	
QY	1261 TTATAAGGGAGCAGCATTTGAACTGGTACCATTTGGTGCAGAAAGAAATATGTCCTGG 1320	
DB	1261 TTATAAGGGAGCAGCATTTGAACTGGTACCATTTGGTGCAGAAAGAAATATGTCCTGG 1320	
QY	1321 CATTAATTTGAGTATGTCATTAATTAATCTTATTAATCTTATTAATCTTATTAAT 1380	
DB	1321 CATTAATTTGAGTATGTCATTAATTAATCTTATTAATCTTATTAATCTTATTAAT 1380	
QY	1381 TTGGAACTGGCGATGGAATTAACCTCAACACCTTATGATGATGAGCTTATGGCGG 1440	
DB	1381 TTGGAACTGGCGATGGAATTAACCTCAACACCTTATGATGATGAGCTTATGGCGG 1440	
QY	1441 TGCTCTCAGGAAAAAATAGATCTTAACTTGAATCTTAACTTCAATATCAAGTTAGCTTAGG 1500	
DB	1441 TGCTCTCAGGAAAAAATAGATCTTAACTTGAATCTTAACTTCAATATCAAGTTAGCTTAGG 1500	
QY	1501 CTCAAATATTTCTTGATTAACATAGGAGGTTGAAATATATATAATAATACTTTAAATTAACG 1560	
DB	1501 CTCAAATATTTCTTGATTAACATAGGAGGTTGAAATATATATAATAATACTTTAAATTAACG 1560	
QY	1561 ATGTTCTTAATGTTTGGGTTAGTTAATAGGTTTCCACCGATCATATAGTAGCTCT 1620	
DB	1561 ATGTTCTTAATGTTTGGGTTAGTTAATAGGTTTCCACCGATCATATAGTAGCTCT 1620	
QY	1621 TCTTTGATGGATGGGTTAGATTATAATGAGTTGCTGGGTTGGATTTTAGATGGGTTAAAT 1680	
DB	1621 TCTTTGATGGATGGGTTAGATTATAATGAGTTGCTGGGTTGGATTTTAGATGGGTTAAAT 1680	
QY	1681 GATTTGGATGGATAATAATAAATGAAATGTTTCTTTTCAATCCGAAAAA 1733	
DB	1681 GATTTGGATGGATAATAATAAATGAAATGTTTCTTTTCAATCCGAAAAA 1733	
RESULT 2		
ID	ADN10568 standard; cdNA; 1538 BP.	
XX	ADN10568	
AC	ADN10568;	

XX	15-JUL-2004	(first entry)	
DT			
XX			
DE	Nicotiana p450 enzyme encoding cDNA SEQ ID NO:153.		
XX			
XX	plant; p450; enzyme; Nicotiana; plant growth regulant; gene therapy;		
KW	phenotype; gene; ss.		
XX			
OS	Nicotiana sp.		
XX			
PN	WO2004035745-A2.		
XX			
PD	29-APR-2004.		
XX			
PF	16-OCT-2003; 2003WO-US032722.		
XX			
PR	16-OCT-2002; 2002US-0418933P.		
XX			
PR	08-JUL-2003; 2003US-0485368P.		
XX			
PR	18-SEP-2003; 2003US-0503989P.		
XX			
PA	(USSM-) US SMOKELESS TOBACCO CO.		
XX			
PI	Xu D;		
XX			
DR	WPI; 2004-348441/32.		
DR	P-FSDB; ADN10569.		
XX			
PT	New isolated nucleic acid molecules and encoded cytochrome P450 enzymes		
PT	from Nicotiana plants, useful for altering plant phenotypes.		
XX			
PS	Claim 1; SEQ ID NO 153; 198pp; English.		
XX			
CC	The present sequence represents a plant p450 enzyme nucleic acid molecule		
CC	isolated from Nicotiana. Also described: (i) an isolated p450 protein		
CC	from Nicotiana; (2) a transgenic plant comprising the nucleic acid		
CC	molecule described above; (3) a method of producing the transgenic plant,		
CC	comprising: (i) operably linking the above nucleic acid molecule with a		
CC	promoter functional in the plant to create a plant transformational		
CC	vector; (ii) transforming the plant with the vector; (iii) selecting a		
CC	plant cell transformed with the transformation vector; and (iv)		
CC	regenerating a transformation plant from the transformed plant cell; and		
CC	(4) a method of selecting a plant containing the above nucleic acid		
CC	molecule, where the plant is analysed for the presence of the above		
CC	nucleic acid sequences. The p450 sequences have plant growth regulant		
CC	activity, and can be used in gene therapy. Compositions and methods from		
CC	the present invention are useful for altering plant phenotypes.		
XX			
SQ	Sequence 1538 BP; 464 A; 278 C; 334 G; 462 T; 0 U; 0 Other;		
	Query Match	23.7%;	Score 410; DB 12; Length 1538;
	Best Local Similarity	58.1%;	Fred. NO. 1.5e-93;
	Matches	816; Conservative	0; Mismatches 555; Indels 33; Gaps 4
Qy	113	CCTCCAGGCGCCATGGAACTTTCCCTATCATAGTATATCTTCCTCATTTATTACTCACTTCT	172
Db	129	CCACCAGTCCATGAACTACCAATACATAGGAAGTATGCTTCATATGGT-----TGGT	182
Qy	173	GATCTAGGCCATGAACGTTTGTAGAGCCCTTGGCTCAAAATTTATGGACCTGTGTATGAGTCTT	232
Db	183	GGACTACCACACCATGTGCTCTTAGAGATTTAGGCCAAAAAATATGGACCGCTTATGCACCTT	242
Qy	233	CAAAATGGCCAAAGTTTCAGCTGTTGTCTATTTCTTCAGCTGAAGCAGCCAAAGAGGTTATG	292
Db	243	CAATTAGTGAAGTTTCTGCAGTTGTGGTTACTTCTCTGATATGGCAAAAGAGTACTA	302
Qy	293	AAAAACTCAGGCTGATGCTCTCGCCCAACGCCCTATGCTGTCTGGAGCGCAGATTTGTGTTT	352
Db	303	AAAACTCATGACATCGCTTTTCGGGTCTAGGCCTAGCCTTTTGGCCCGCGAGATTGCTGT	362
Qy	353	TATTAATCGGAAAGATGCTTGTGTTGCTTCATATGGAGATCACTGGAGCGAGATGAAGAA	412
Db	363	TACATAGGCTGATCTTGGCGTTTGGCCCTATGGCGATTATTTGAGACAATGGGTAAA	422

Db	1503	TACTTGGTTGGCACTCCTTATCAA	1526	Db	303	AAAACTCATGACATCGCTTTTCGCGCTAGCGCTAGCTTTTGGCCCCGGAGATGTCGTCT	362
RESULT 3				Qy	353	TATAATCGGAAAGATGCTTGTGTTTCTTATATATGAGATCACTGGAGCGCATGAAGAAA	412
ADZ65256				Db	363	TACAATAGGTCGTGATCTGCGTTTGGCCCTATGGCGATTATGGAGACAAATGCGTAAA	422
ID ADZ65256 standard; cDNA; 1538 BP.				Qy	413	ATTTGGATACCTTGAATTTCTGAGTCCGAAAGATTCATCTCCAGGTAAATCCGAGAG	472
AC ADZ65256;				Db	423	ATATGTGTCTTTGGAAAGTCTCAGTGCCCAAGAAATGTTGGACATATAGCTCTATTAGGGCG	482
XX				Qy	473	GAAGAAATGGAGGATGCCATCACATTCCTCCGTTTCGAAAGCCGGATCTCCGGTCAATATT	532
DT 14-JUL-2005 (first entry)				Db	483	GATGAAGTCTTCTGCTCTTAAATTTATCCGGTCATCTTCTGGTGAGCCCTGTTAAATT	542
XX Tobacco full length cDNA for cytochrome p450 clone D90A-BB3.				Qy	533	ACAAAGATCAATTTATGGCATTTAATTTTCGATCATGATAAGAAATCCGTTGGTAA---T	589
XX Cytochrome p450; ss; secondary metabolite; ethylene; plant; senescence;				Db	543	ACGGAAGGATCTTTTGTTCACAAGCTCCATGACATGTAGATCAGCGTTGGGCAAGTA	602
XX normicotine; transgenic plant; gene.				Qy	590	TGTAAGCAAAAGAAAGATTGCTGAGTGTGCCGATGCAGTCMAATGAGCGACGACGAGT	649
XX Nicotiana tabacum.				Db	603	TTCAAGGAGCAAGCAAAATTTATACAACTAATTAAGAAAGTTATATCTCTTAGCAGAGGG	662
XX WO2005038033-A2.				Qy	650	TTTGGCACCGCAGACGCTTTTCCGACGTGGAAATTTACTTTCACATATATCATTTGGAGCTGAG	709
XX 28-APR-2005.				Db	663	TTTGATGGCTGACATATTCCTTTCATCAAGTCTCTTCATGTGCTCAGTGGATGAAG	722
XX 15-OCT-2004; 2004WO-US034065.				Qy	710	TCAAAACCCAGCGTTTGCATCAGGAGATTGACGATATATTGGAAGATTTCTTAATGAA	769
XX 16-OCT-2003; 2003US-00686947.				Db	723	GGTAAGATTATGAATGCACACATGAAGGTAGTCTATTGTTGAGAAATGTATCAACGAG	782
XX 29-APR-2004; 2004US-0566235P.				Qy	770	CAC-----AAAGCCAAATAGCCCTTTTGAAGCGGATAACTTAATG	808
XX 03-SEP-2004; 2004US-00934944.				Db	783	CACAAGAAAAATCTTGGCAATTTGGGAAAACTAATGGAGCGTTAGGAGGTGAAGATTAAIT	842
XX (USNM-) US SMOKELESS TOBACCO CO.				Qy	809	GATGTTCTATTGAATCTTCAAAAAAATGGAACCGTTCCAGTGCCAGTGACAAACGAAAGC	868
XX Xu D;				Db	843	GATGTTCTTCTTAAACTTATGAATGATGAGGCGCTTCAATTTCTTATCACCACGACAAC	902
XX WPI: 2005-315717/32.				Qy	869	ATCAAGCATCCGTTTTCGAAATGTTTACTGCGGGAGCGGAAACAACTTCGAAAGCTACA	928
XX P-PSDB; ADZ65257.				Db	903	ATCAAGCTATAATCTTTGACATGTTGCTGCTGGAACAGAGACTTCATCGTCAACAAAT	962
XX New nucleic acid molecule encoding cytochrome P450 enzymes in Nicotiana,				Qy	929	GAATGGGTAAATGGGAGAGCTGATGAAAAATCCAACTGAATTAAGAAAAAGCACAAAGAA	988
XX useful in developing tobacco plants with altered phenotypes.				Db	963	GTGTTGGCTATGGTGGAAATGGTGAATAATCAACTGTATTTGCGAAAGCTCAGCAGAA	1022
XX Disclosure; SEQ ID NO 153; 226pp; English.				Qy	989	GTTAGACAAGTATTTGGTGAATGGGAAAGTGTGATGAATCAAGATTTTCATGATTTGAAA	1048
XX The invention relates to an isolated nucleic acid molecule (I) from				Db	1023	GTAAGAGATGCATTTAGAGAAAAAGAAACTTTTGATGAAATGATGTGGAGGAGCTAAAC	1082
XX Nicotiana, where the nucleic acid molecule comprising any of the 59				Qy	1049	TTCTTCAAAGTTAGTGTAAAGAAACTCTAAGATTACATCTCTCCGCTT---GTCTTGATT	1105
XX nucleic acid sequences of SEQ ID NOS: 299-357 (NOTE:the claims refer to				Db	1083	TATCTAAAGTTAGTCAATTAAGAAACTCTAAGACTTCTCCACCGGTTCCACTTTTGCTC	1142
XX SEQ ID NOS 299-357 as nucleic acids but these sequences (apart from SEQ				Qy	1106	CCGAGGAGTGTAGAGAAACACGAATTTGATGATATGAATTTTCATCCGAACACTCGA	1165
XX ID NO 356) are all proteins and appear as ADZ65402-ADZ65460. The nucleic				Db	1143	CAAAGAGATGTAGGGAAGAGACAAATATAACGGCTACACTATTTCTGCTGAAGACCAA	1202
XX acids of the invention encode cytochrome P450 enzymes whose expression is				Qy	1166	ATTGTTGTAATGCTTTGGGCGATAGGAAGAGATCTTAATATCTTGGTCGGAACCTGGAAG	1225
XX induced by ethylene and/or plant senescence. Also included are a				Db	1203	GTCATGGTTAATGTTTGGGATTTGGGAAGAGATCCAAATATTTGGGATGATCAGAAACT	1262
XX transgenic plant comprising (I), a method of producing a transgenic				Qy	1226	TTTAACCCAGAAAGGTTTAAAGATTGTGCAATTTGATTATAAGGAGCGACATTTGNACTG	1285
XX plant, a method of selecting a plant containing a nucleic acid molecule				Db	1263	TTTAAGCCAGAGAGATTGAGCAGTGTCTAAGGATTTTGTGTTGTAATTAATTTGTAATAT	1322
XX (where the plant is analyzed for the presence of nucleic acid sequence of				Qy	1286	GTACCATTTGGTGAGGAAAAAGAAATATGTCCTCGCATTTACTTTCAGCTATTACCAATTTG	1345
XX ADZ65402-				Db	1323	CTTCCATTTGGTGGTGGAGGAGATTGTCCAGGGATTTGTTGGTTTGTAGCTAATGCT	1382
XX Sequence 1538 BP; 464 A; 278 C; 334 G; 462 T; 0 U; 0 Other;				Qy	1346	GAGTATGTCAATTATAAATCTATTATATCAATTTTAATTTGGAACTCGGCCGATGAATACA	1405
Query Match 23.7%; Score 410; DB 14; Length 1538;				Db	1383	TATTTGGCCATTGGCTCAATTACTTTATCACTTTGATTGGGAACTCCCACTGGAATCAAA	1442
Best Local Similarity 58.1%; Pred. No. 1.5e-93;							
Matches 816; Conservative 0; Mismatches 555; Indels 33; Gaps 4;							
Qy 113	CTCTCAGGCGCATGGAAGTTTCTATCATAGTAAATCTCTCTCATTTATTACTCACTTCT	172					
Db 129	CCACAGGTGCTTGGAACTACCAATACCTAGGAAGTATGCTTCATATGTT-----TGGT	182					
Qy 173	GATCTAGGCGCATGAAAGCTTTTAGAGCTTTGGCTCAAAATTTATGGACCTCTTATGATCTT	232					
Db 183	GGACTACCAACCATGCTCTTAGAGATTTAGCCAAAAATATAGGACCGCTTATGCACCTT	242					
Qy 233	CAAAATGGCCAAAGTTTCACTGTTGTCATTTCTCAGCTGGAAGCGCCAAAGAGTTATG	292					
Db 243	CAATAGGTGAAGTTTCTGCAGTTGTGGTTACTTCTCTGATATGGCAAAAGAACTACTA	302					
Qy 293	AAAACCTCAGCTGATGCTCTCGCCCAACCGCCCTATGCTTGTGGACGACAGATTGTTT	352					

Db 1143 CCAAGAGAAATGTAGGAAGAGACAATAATATAAAGCGCTACATATTCCTGTTAAAGCCAAA 1202
Qy 1166 ATTGTTGTGAATGCTTTGGCGCATAGGAAGAGATCCCTAATCTATCTGTCGGAACCTGGAAG 1225
Db 1203 GTCATGTTAATGTTTGGCGATTGGAGAGATCCAAAATATTGGATGATCCAGAACT 1262
Qy 1226 TTTAACCCAGAAAGTTTAAAGATTGTGCAATTGATTAATAAGGAGCAGACATTGGAATG 1285
Db 1263 TTTAAGCCAGAGAGATTTGAGCAGTGCTCTAAGGATTTTGTGGTAATAATTTTGAATAT 1322
Qy 1286 GTACCATTTGGTCGGAAGAAAGATATGTCCTGCATCTACTTCAGCTATTACCAATTTG 1345
Db 1323 CTTCCATTGGTGGTGAAGGAGGATTTGTCCAGGAATTCGTTGGTTTACCTAATGCT 1382
Qy 1346 GAGTATGTCATTATAAATCTATTATATCATTTTAAATTTGGGAACCTGCGCATGGAAATTACA 1405
Db 1383 TATTTGCCATTGGCTCAATTACTTTATCATTTTGTATGGGAACCTCCCACTGGATCAA 1442
Qy 1406 CTTCAAAACATTGATATGACTGAAGCTATTGGCGGTGCTCTCAGGAAAAAATAAGATCTT 1465
Db 1443 CCAAGCGACTTGGACTTGACTGAGTTGGTTGGAGTAACCTGCCGCTAGAAAAAAGTGACCTT 1502
Qy 1466 AGTTGATTCCTATTCATCAATCAA 1489
Db 1503 TACTTGGTGGCACTTCCTTATCAA 1526

RESULT 5

AEE06775

ID AEE06775 standard; cDNA; 1538 BP.

AC AEE06775;

XX

DT 26-JAN-2006 (first entry)

XX

DE Tobacco p450 cDNA SEQ ID NO: 162.

XX

KW Plant breeding; plant; ss; coding sequence; cytochrome p450;

KW transgenic plant; senescence.

XX Nicotiana tabacum.

XX

PN WO2005111217-A2.

XX

PD 24-NOV-2005.

XX

PF 27-APR-2005; 2005WO-US014803.

XX

PR 29-APR-2004; 2004US-0566235P.

PR

PR 03-SEP-2004; 2004US-00934944.

PR

PR 03-SEP-2004; 2004US-0607357P.

PR

PR 17-SEP-2004; 2004US-00943507.

PR

PR 15-OCT-2004; 2004WO-US034065.

PR

PR 15-OCT-2004; 2004WO-US034218.

PR

PR 25-JAN-2005; 2005US-0646764P.

PR

PR 24-MAR-2005; 2005US-0665097P.

PR

XX (USSN-) US SMOKELESS TOBACCO CO.

XX

XX Xu D, Nielsen MT;

XX

XX WPI; 2005-786788/80.

XX

XX P-PDB; ABE06776.

XX

XX Producing a tobacco plant having decreased expression of a nicotine

XX demethylase gene comprises crossing a first tobacco plant with a second

XX tobacco plant and germinating the collected seed of an F1 progeny plant.

XX

XX Claim 32; SEQ ID NO 162; 641pp; English.

XX

XX The invention relates to a breeding method for producing a tobacco plant

XX with reduced expression of a nicotine demethylase gene comprises crossing

CC a first tobacco plant with variant nicotine demethylase gene expression
CC with a second tobacco plant with at least one phenotypic trait to produce
CC an F1 progeny plant, the seed of which is collected and germinated to
CC produce a tobacco plant having reduced expression of a nicotine
CC demethylase gene. Also included are breeding a nicotine demethylase
CC deficiency trait into a tobacco plant, producing a tobacco seed
CC developing a tobacco plant in a tobacco breeding program, a tissue
CC culture of regenerable tobacco cells obtained from the tobacco plant of
CC the invention, producing a tobacco product, a breeding method for
CC producing a tobacco plant having a modified attribute, a method of
CC breeding an attribute into a tobacco plant, a tobacco plant or its
CC components produced by the method of breeding a nicotine demethylase
CC deficiency trait into a tobacco plant, producing tobacco seed, producing
CC a tobacco plant having a modified attribute or developing a tobacco plant
CC in a tobacco breeding program, an isolated genetic marker comprising a
CC nucleic acid sequence that is substantially identical to a nucleic acid
CC sequence given in the specification (the nucleic acids comprise isolated
CC cytochrome p450 cDNAs), an expression vector comprising the nucleic acid
CC sequence, a plant or plant component comprising the isolated nucleic acid
CC sequence, a plant produced from a germinated seed of the plant, reducing
CC the expression or enzymatic activity of a constitutive, or an ethylene
CC induced or senescence induced tobacco polypeptide in a plant cell, and
CC increasing the expression or enzymatic activity of a constitutive, or an
CC ethylene or senescence induced tobacco polypeptide in a plant cell. The
CC phenotypic trait comprises disease resistance, high yield, high grade
CC index, curability, curing quality, mechanical harvestability, holding
CC ability, leaf quality, height, maturation, stalk size, or leaf number per
CC plant. The breeding method for producing a tobacco plant having decreased
CC expression of a nicotine demethylase gene is useful developing desirable
CC (non-genetically engineered) germplasm. The plant is useful in producing
CC (smokeless) tobacco products. The tobacco product is a moist or dry
CC snuff, a chewing tobacco, a cigarette product, a cigar product, a
CC cigarillo, a pipe tobacco, or bidis. The p450 cDNAs were isolated using
CC degenerate PCR primers designed against cytochrome p450 motifs. The
CC present sequence is a cytochrome p450 cDNA of the invention.
XX Sequence 1538 BP; 464 A; 278 C; 334 G; 462 T; 0 U; 0 Other;

Query Match 23.7%; Score 410; DB 14; Length 1538;

Best Local Similarity 58.1%; Pred. No. 1.5e-93;

Matches 816; Conservative 0; Mismatches 555; Indels 33; Gaps 4;

Qy 113 CCTCAGGCGCATGGAGTTTCCTATCATAGTAATCTCTCTCATTTTATTACTCATTCT 172
Db 129 CCACCGAGTCCATGGAAACTACCAATACCTAGAGATGCTTCATATGGT-----TGGT 182
Qy 173 GATCTAGGCCATGAACGTTTTAGAGCCTTGGCTCAAATTTATGGACCTGTTATGAGTCT 232
Db 183 GGACTACCACACCATGCTCTTAGAGATTTAGCCAAAAAATATGGACCGCTTATGCACCTT 242
Qy 233 CAATTTGCCCAAGTTTCAGCTGTTGTGTCATTTCTTCAGCTGAAGCAGCCAAAGAGGTTATG 292
Db 243 CAATTTAGGTGAAGTTTCTGTCAGTTGTGGTTACTTCTCTCGATATAGCAAAAGAACTACTA 302
Qy 293 AAAAATCAGGCTGATGCGCTTCGCCCAACGCTATCGTCTTGGAGCCACAGATTGTGTTT 352
Db 303 AAAAATCATGATCGCTTTCCGCTCTAGGCTAGCCTTTTGGCCCCCGAGAGATTGCTGT 362
Qy 353 TATAATCGAAAGATGCTCTTGTGTTCTTATATGGAGATCACTGGAGGCAGATGAAGAAA 412
Db 363 TACAATAGTCTGATCTTTCGCTTTTGGCCCTATGCGATTTATGGAGACAAATGCGTAAA 422
Qy 413 ATTTGGATACATTGAATTTCTGAGTGCCAAAAAAGTTCAATCTCCAGGTTAATCCGAGAG 472
Db 423 ATATGTGCTTGTGAAGTGTCTCAGTSCCAAGAAATGTTCCGACATATAGCTCTATTATGGCGC 482
Qy 473 GAAGAAATGGAGGATGCCATCATTCTCCGTTCCGTTCCGTTCCGTTCCGTTCCGTTCCGTT 532
Db 483 GATGAAGTCTTCTGCTCTCTTAATTTTATCCGGTCATCTTCTGTTGGAGCTCTTAAATTT 542
Qy 533 ACAAGATCATTTATGGCATTATAATTTCCGATCATGATAAGAACATCCGTTGGTAA---T 589
Db 543 ACGGAAAGGATCTTTTGTTCACAAGCTCCATGACATGTAGATCAGCGTTTGGGCAAGTA 602

Qy	590	TGTAAGCAAAAAGAAAGATTGCTGAGTGTTCGCCGATGCAGTCAATAGGCGCAGCCAGCAGT	649
Db	603	TTCAAGGAGCAAGCAAAATTTATACAACTAATTAAGAAGTTATACTCTTAGCAGGAGGG	662
Qy	650	TTTGGCACCGCAGACGCTTTTCCGACGCGGAAATTACTTCTACTATATCATATTGGAGCTGAG	709
Db	663	TTTGATGTGGCTGCACATATCCCTTTCATACAAGTCTCTTCATGTGCTCAGTGGSAATGAAG	722
Qy	710	TCAAAACCCAGCGCTTGGCATCAGGAGATTGACGATATCTTGAAGAGATTCTTAATGAA	769
Db	723	GGTAAGATTATGAATGCACACCATTAAGTAGATGCTATTGTTGAGAAATGTCATCAACGAG	782
Qy	770	CAC-----AAAGCCAAATTAAGCCTTTTGAAGCGGATACTTTAATG	808
Db	783	CACAAGAAAAATCTTGGCAATTTGGGAAAACTAATGGAGCGTTAGGAGGTGAAGATTTAATT	842
Qy	809	GATGTTCTATTGAATCTTCAAAAAAATGGAACGTTCCAGTGCAGTGCACAAACGGAAGC	868
Db	843	GATGTTCTTTAAACCTTATGAATGATGGAGGCCCTTCAATTTCTATCACCACAGCAAC	902
Qy	869	ATCAAAAGCATCCGTTTTCGCAATGTTTACTGCCGGAGCGAAAACTTCGAAAGCTACA	928
Db	903	ATCAAGCTATAATCTTTGACATGTTTCTGCTGGAACAGAGACTTCATCGTCAACATTT	962
Qy	929	GAATGGGTAAATGGCAGAGCTGATGAAAAATCCAACTGAACTAAGAAAGCACAAGAGAA	988
Db	963	GTGTGGGCTATGGTGGAAATGGTGAAAAATCCAACTGTATTTGCGAAAGCTCAAGCAGAA	1022
Qy	989	GTTTAGACAAGTATTTGTTGTAATCGGAATCGGAATGTCATGAATCAAGATTCATGATTGAAA	1048
Db	1023	GTAAGAGATGCATTTAGAGAAAAAGAACTTTTGTGATAAAATGATGTGGAGGAGCTAAAC	1082
Qy	1049	TTCTTCTCAAGTTAGTGGTTAAAGAACTCTAAGATTTACATCCTCCGGTT---GTCTTGATT	1105
Db	1083	TATCTAAGTTAGTCAATTAAGAATACTTAAGACTTCATCCACGGTTCCACTTTTGCTC	1142
Qy	1106	CCGAGGAGGTGTAGAGAAACAACAGCAATGTATGGATATGAAATTCATCGAACACTCGA	1165
Db	1143	CCAAGAAATGTAGGGAAGACAAAAATATAAACGGCTACATACTCTCTGTAAGACCAA	1202
Qy	1166	ATTGTTGTGAATGCTCGGGCATAGGAGAGATCCTAATACTTTGGTCCGAACTGGGAAG	1225
Db	1203	GTCAATGTTAATGTTTGGCAATTGGGAAGAGATCCAAAAATTGGGAATGATGCAAGAACT	1262
Qy	1226	TTTAAACCCAGAAAGGTTTAAAGATTGTGCAATTGATTATAAAGGACGACATTTTGAAC	1285
Db	1263	TTTAAAGCCAGAGATTTGACGAGTCTCTAAGGATTTTGTGTGTAATATTTTGAATAT	1322
Qy	1286	GTACCATTTGGTGAGAAAAAGAAATATGCTCGGCAATTACTTCAGCTATTAACAATTTG	1345
Db	1323	CTTCCATTTGGTGGTGAAGAGGATTTGTCCAGGGATTTTCGTTGGTTAGCTTAATGCT	1382
Qy	1346	GAGTATGTCATTATAATCTATTATATCATTTTAATGGGAACTGGCCGATGGAAATTACA	1405
Db	1383	TATTTGGCAATGGCTCAATTACTTTATCACTTTGATTTGGGAATCTCCCACTGGGAATCAA	1442
Qy	1406	CCTCAAAACACTTGGATATGACTGAAGCTATTGGCGGTGCTCTCAGAAAAAAAATAGATCTT	1465
Db	1443	CCAAGCGACTTGGACTTGACTGAGTTGGTTGGAGTAATCTGCCGCTAGAAAAAGTGACCTT	1502
Qy	1466	AAGTTGATTCCTATTTCATATCAA	1489
Db	1503	TACTTGGTTGCACTCCTTATCAA	1526

RESULT 6

REF ID: AEF97168

ID AEF97168 standard; cDNA: 1538 bp.

XX

AC AEF97168;

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

Tobacco cytochrome P450 cDNA clone D90A-BB3, SEQ ID NO:162.

Plant breeding; crop improvement; secondary metabolite; genetic marker; plant; cytochrome P450; CYP gene; ss.

Nicotiana tabacum.

US2006041949-A1.

23-FEB-2006.

27-APR-2005; 2005US-00116881.

13-NOV-2001: 2001US-0337684P

11-JAN-2002; 2002US-0347444P

12-MAR-2002; 2002US-0363684P

16-OCT-2002; 2002US-0418933P

13-NOV-2002; 2002US-00293252

10-JAN-2003; 2003US-00340861
13-MAR-2003; 2003US-00397346

12=MAR=2003; 2003US=00387346
08=JUL=2003; 2003US=04853689

08-JUL-2003; 2003US-0483368F
18-SEP-2003; 2003US-0503989P

16-OCT-2003: 2003US-00686947

29-APR-2004; 2004US-0566235P.

03-SEP-2004; 2004US-00934944

03-SEP-2004; 2004US-0607357P.

17-SEP-2004; 2004US-00943507.

15-OCT-2004; 2004WO-US034065;
15-OCT-2004; 2004WO-US034218

15=UCI=2004; 2004WO=05034218;
25=JAN=2005; 2005IIS=0646764P

23-JAN-2005; 2005US-0846784E
24-MAR-2005: 2005US-0665097P

24-MAR-2005; 2005US-0665451P

19-APR-2005; 2005US-00110062.

(USSM-) US SMOKELESS TOBACCO

Xu D, Nielsen MT;

WPI; 2006-182895/1

P-PSDB; AEF97169.

New breeding methods

decreased expression of a nicotine demethylase gene comprises crossing germinating F1 progeny seed to produce the tobacco plant.

Claim 63: SEO TD NO 162: 511pp. English

The invention relates to a breeding method for producing a tobacco plant having decreased expression of a nicotine demethylase gene. The method involves crossing a first tobacco plant having variant nicotine demethylase gene expression with a second tobacco plant containing at least one phenotypic trait (e.g., disease resistance, high yield etc.) to produce an F1 progeny plant; collecting the seed of the F1 progeny; and germinating the seed to produce a tobacco plant having decreased nicotine demethylase expression. The invention also relates to a tobacco plant or its components produced using the method of the invention; a tissue culture of regenerable tobacco cells obtained from such plants; a tobacco product produced from such plants; and a method of breeding a tobacco plant with a modified attribute comprising variant expression of a cytochrome P450 polynucleotide (including the nicotine demethylase gene). The invention further relates to isolated constitutive, ethylene-induced or senescence-induced genetic markers comprising tobacco cytochrome P450 or nucleic acid molecules, including nicotine demethylase sequences; methods for reducing or increasing the expression or activity of polypeptides encoded by these nucleic acid molecules; expression vectors, plants or plant components comprising one of these polynucleotides; and tobacco products produced from such plants. The methods of the invention are useful for the breeding (especially marker assisted breeding) of tobacco plants with decreased nicotine demethylase expression or with altered cytochrome P450 expression or activity. Altered expression of such enzymes can result in a change in the composition of secondary metabolites such as alkaloids (e.g., nicotine), phenylpropanoids,

CC terpenoids, lipids, cyanogenic glycosides and glucosinolates, with
CC effects on the flavor or aroma of plant products. They may also affect
CC herbicide tolerance, resistance to disease or insects, quality factors
CC related to undesirable constituents, structural traits, fiber content,
CC leaf yield, ripening, leaf curing or storage properties. Tobacco plants
CC of the invention in which the expression of cytochrome P450 genes is
CC altered may have desirable traits such as altered levels of nicotine
CC or N'-nitrosonicotine. Such plants can be used in the production of
CC tobacco products such as moist or dry snuff, chewing tobacco, cigarettes,
CC cigars, cigarillos, pipe tobacco, bidis or smokeless tobacco products.
CC The present sequence represents a specifically claimed tobacco cytochrome
CC P450 polynucleotide useful in marker assisted breeding methods. Note: The
CC sequence data for the US patent is also available in electronic format
CC directly from the US patent office at
CC seqdata.uspto.gov/sequence.html?DocID=20060041949.

XX
SQ Sequence 1538 BP; 464 A; 278 C; 334 G; 462 T; 0 U; 0 Other;

Query Match 23.7%; Score 410; DB 15; Length 1538;
Best Local Similarity 58.1%; Pred. No. 1.5e-93;
Matches 816; Conservative 0; Mismatches 555; Indels 33; Gaps 4;

QY	113	CCTCAGGCGCATGGAAGTTTCTCATCATAGTAAATCTCTCATTTTACTCACTTCT	172
DB	129	CCACAGGTCATGGAACTCAATACCTAGGAAGTATGCTTCATATGGT-----TGGT	182
QY	173	GATCTAGGCGCATGAACGTTTTAGAGCCTTTGGCTCAAAATTTATGGACCTGTTATGAGTCTT	232
DB	183	GGACTACCAACACATGCTCTTAGAGATTTAGCCAAATAATATGGACCGCTTATGACCTT	242
QY	233	CAAAATGGGCAAGTTTTCAGCTGTTGTCATTTCTTCAGCTGAAGCAGCCAAAGAGTTATG	292
DB	243	CAATTAGGTGAAGTTTCTGCAGTTGTTGTTTACTTCTCTGATATGGCAAAAGAACTACTA	302
QY	293	AAAACTCAGGCTGATGCGCTTCGCCAACGCCCTATGCTCTGCAGCAGACAGATTGTTT	352
DB	303	AAAACTCATGACATGCTTTCGCGCTAGGCTTAGGCTTTGGCCCGGAGATGTTCTGT	362
QY	353	TATAATCGGAAAGATGTTGTTGTTTTCATATGGAGATCACTGGAGCGAGATGAAGAAA	412
DB	363	TACAAATAGTCTGATCTTGGCTTTGGCCCTATGCGGATTTAGGACAAATGCGTAAA	422
QY	413	ATTGGAATCTGAATTTCTGAGTGCCAAAGTTCAATCTTCAGGTTATTCAGAG	472
DB	423	ATATGTGTTTGGAAAGTGCTCAGTGCCAAAGATGTTTCGGACATATAGCTCTTATAGGCGC	482
QY	473	GAAAGAAATGGAGATGCCATCACTTCTCCGTTTCGAAAGCCGATCTCCGCTCAATATT	532
DB	483	GATGAAGTTCTTCGTTCTCTTAATTTATCCGGTCACTTCTGTTGAGGCTGTTAATATT	542
QY	533	ACAAAGATCATTTATGGCATTTATAATTTTCGATCATGATAAGAACATCCGTTGGTAA---T	589
DB	543	ACGMAAAGGATCTTTTGTTCACAGCTCCATGACATGTAGATCAGCGTTTGGGCAAGTA	602
QY	590	TGTAGCAAAAAGAAAGATTGCTGAGTGTGCGGATGTCAGTCAATGAGCGACGCGAGT	649
DB	603	TTCAAGGACCAAGACAAATTTATACAACTAATTAAGAAGTTTATCTCTTAGCAGAGGG	662
QY	650	TTTGGCAGCGACAGCCTTTTCGAGCTGGAATTTACTTCACTATATATCTGGAGCTGAG	709
DB	663	TTTGATGTGCTGACATATTCCTTTCATCAAGTCTCTTCATGTGCTCAGTGGATGAAG	722
QY	710	TCAAAACCCAGGCGTTTGCATCAGGAGATTGACGATATATCTTGAAGAGATTCTTAAATGA	769
DB	723	GGTAAAGATTATGAATGCAACACATAGGTAGTGTCTTATGTTGAGAATGTCTATCAACGAG	782
QY	770	CAC-----AAAGCCATAAGCCTTTTGAAGCGGATCACTTAATG	808
DB	783	CACAAGAAAAATCTTCCAAATGGGAAAACTAATGGAGCGTTAGGAGGTGAAGATTTAAT	842
QY	809	GATGTTCTATTGAATCTTCAAAAAATGGAACGTTCCAGTGCAGTGACAAACGAAAGC	868
DB	843	GATGTTCTTCTAAACCTTATGAATGATGAGGCGCTTCAATTTCTTATCAACCAACGACAC	902

QY	869	ATCAAGACATCCGTTTGTGCAAAATGTTTACTCCCGGAGCGAAACAACTTCGAAAGCTACA	928
DB	903	ATCAAGCTATAATCTTTGACATGTTTGTCTCTGGAACAGAGACTTTCATCGTCAACAATT	962
QY	929	GAATGGGTAAATGGCAGAGCTCATGAAAAATCCAACTGAACTAAGAAAAAGCAAGAAGAA	988
DB	963	GTGTGGGCTATGGTGAATGGTGAATAATCAACTGTATTTTGGAAAGCTCAGACAGAA	1022
QY	989	GTTAGACAAGTATTTGGTGAATGGAAAAAGTTGATGAATCAAGATTTTCATGATTTGAAA	1048
DB	1023	GTAAGAGATGCATTTAGAGAAAAAGAACTTTTGTATGAAAAATGATGTGGAGGAGCTAAAC	1082
QY	1049	TTCTTCAAGTTAGTGGTTTAAAGAACTCTAAGATTACATCTCCGGTT---GTCTTTGATT	1105
DB	1083	TATCTAAAGTTAGTCATTTAAAGAAACTCTAAGACTTTCATCCACCGGTTCCACTTTTGCTC	1142
QY	1106	CCGAGGAGTGTAGAGAAACAAACAGAAATTCATGGATATGAAATTCATCCGAACACTCGA	1165
DB	1143	CAAAGAAATGTAGGAAAGAGACAAATATAAACGGCTACACTATTCCTGTTAAGACCAAA	1202
QY	1166	ATTGTTGTGAATGCTTGGCGGATAGGAAGAGATCCATACTTGTGTCGGAACCTGGAAAG	1225
DB	1203	GTCAATGTTAATGTTTGGGCATTTGGGAAGAGATCCAAAAATATTGGGATGATGCAGAACT	1262
QY	1226	TTTAACCCAGAAAGGTTTAAAGATTGTCCAATTGATTAATAAGGACGACATTTGAACTG	1285
DB	1263	TTTAAGCCAGAGAGATTTTGGACAGTCTCTAAGGATTTTGTGGTAAATTTTGAATAT	1322
QY	1286	GTACCATTTGTGCGAGGAAAAAGAAATATGTCTCGCATTTACTTTCAGCTATTATCAAAATTTG	1345
DB	1323	CTTCCATTTGTTGGTGGAAAGAGGATTTGTCCAGGGATTTTCGTTTGGTTAGCTAATGCT	1382
QY	1346	GAGTATGTCAATTAATAATCTTATATATCATTTTAAATGGGAACCTGGCCGATGGAATTACA	1405
DB	1383	TATTTGCCATTTGGCTCAATTTACTTTATCATTTGATGGGAACCTCCCACTGGAATCAAA	1442
QY	1406	CCTCAACACATTGATATGACTGAAGCTATTGGCGGTGCTCTCAGGAAAAAATAGATCTT	1465
DB	1443	CCAAAGCGATTTGGACTTGCATGAGTTGGTGGAGTAACTGCCGCTAGAAAAAAGTGACCTT	1502
QY	1466	AAGTTGATTCTTATTCATATCAAA	1489
DB	1503	TACTTGGTTGGACTCTTATCAAA	1526

RESULT 7
ADN10646

ID ADN10646 standard; cDNA; 1673 BP.

XX ADN10646;

XX 15-JUL-2004 (first entry)

XX Nicotiana p450 enzyme encoding cDNA SEQ ID NO:231.

XX plant; p450; enzyme; Nicotiana; plant growth regulant; gene therapy;
phenotype; gene; ss.

XX Nicotiana sp.

XX WO2004035745-A2.

XX 29-APR-2004.

XX 16-OCT-2003; 2003WO-US032722.

XX 16-OCT-2002; 2002US-0418933P.

XX 08-JUL-2003; 2003US-0485368P.

XX 18-SEP-2003; 2003US-0503989P.

XX (USNM-) US SMOKELESS TOBACCO CO.

XX 16-OCT-2003; 2003WO-US032722.
XX PF
XX PR
XX PR 16-OCT-2002; 2002US-0418933P.
XX PR 08-JUL-2003; 2003US-0485368P.
XX PR 18-SEP-2003; 2003US-0503989P.
XX PA
XX PA (USSM-) US SMOKELESS TOBACCO CO.
XX PI
XX PI Xu D;
XX WPI; 2004-348441/32.
XX DR P-PSDB; ADN10625.
XX PT
XX PT New isolated nucleic acid molecules and encoded cytochrome P450 enzymes
XX PT from Nicotiana glauca, useful for altering plant phenotypes.
XX PT
XX PT Claim 1; SEQ ID NO 209; 199pp; English.
XX PT
XX PT The present sequence represents a plant p450 enzyme nucleic acid molecule
XX CC isolated from Nicotiana glauca. Also described: (i) an isolated p450 protein
XX CC from Nicotiana glauca; (ii) a transgenic plant comprising the nucleic acid
XX CC molecule described above; (iii) a method of producing the transgenic plant,
XX CC comprising: (i) operably linking the above nucleic acid molecule with a
XX CC promoter functional in the plant to create a plant transformational
XX CC vector; (ii) transforming the plant with the vector; (iii) selecting a
XX CC plant cell transformed with the transformation vector; and (iv)
XX CC regenerating a transformed plant from the transformed plant cell; and
XX CC (4) a method of selecting a plant containing the above nucleic acid
XX CC molecule, where the plant is analysed for the presence of the above
XX CC nucleic acid sequences. The p450 sequences have plant growth regulant
XX CC activity, and can be used in gene therapy. Compositions and methods from
XX CC the present invention are useful for altering plant phenotypes.
XX CC
XX CC Sequence 1673 BP; 493 A; 309 C; 357 G; 514 T; 0 U; 0 Other;
XX CC
XX CC Query Match 23.1%; Score 400.4; DB 12; Length 1673;
XX CC Best Local Similarity 57.7%; Pred. No. 4.4e-91;
XX CC Matches 810; Conservative 0; Mismatches 561; Indels 33; Gaps 4;
XX CC
XX CC 113 CCTCAGGCGCATGGAAGTTTCTCATCATAGTAACTCTCCCTCAATTTATTACTCACTTCT 172
XX CC 135 CCACAGGTCCATGGAACACTCACTACTAGGAAGTATGCTTCATATGTT-----TGGT 188
XX CC 173 GATCTAGGCGCATGGAAGTTTCTAGAGCTTTGGCTCAAAATTTATGGACCTGTTATGAGTCTT 232
XX CC 189 GCACTACACACCATGCTCTTAGAGATTAGCCAAATAATATGGACCATTTATGACCTT 248
XX CC 233 CAAATTTGGCCAAATTTGATGCTTTGATTTCTTCAAGCTGGAAGCGCAAGAGGTTATG 292
XX CC 249 CAATTAGGTGAAGTTTCTGCGTTTGGTTACTTCTCTGATACGCGCAAAAGAGATTTA 308
XX CC 293 AAAAATCAGGCTGATGCTTCCGCCAACCCCTATCGTCTTGGACGCGACAGATGCTGTT 352
XX CC 309 AAAAATCATGATGATGCTTTTGGCTTAGGCTTAGGCTTTTGGCCCCGGAGATTTGCTGT 368
XX CC 353 TATAATCGGAAGAGATGCTTTGTTTGTCTCATATGGAGATCACTGGAGGCGAGATGAAGAA 412
XX CC 369 TACAATAGTCTGATCTAGCTTTTGGCCCTATGGCGACTATGGAGCAATATGCGTAA 428
XX CC 413 ATTTGATATCTTGAATTTCTGATGTCGCAAAAGTTTCAATCTCCAGGTTTAATCCGAGAG 472
XX CC 429 ATATGTGCTTGGAGTGTCTAGTGTGCAAGAAATGTTTGGACATTTAGCTCTATTAGGCGG 488
XX CC 473 GAAGAAATCGGAGATGCCATCATCTTCCCTCGGTGGAGCGGATCTCCGCTCAATATT 532
XX CC 489 AATGAAGTTCTTCGCTCTCATTAATTTTATCCGGTCACTCTCTGGTGAACCTAATATGTT 548
XX CC 533 ACAAGATCATTTATGGCATTTAATTTTCGATCATGATGAAGAACATCCGTTGGTAA---T 589
XX CC 549 ACGGAAGGATCTTTTGTTCACAGCTCCATGACATGTAGATCAGCGTTTGGCGAAGTG 608
XX CC 590 TGTAGCAAAAGAAAGATTGCTAGTGTGGGATGCGAGTCAATGAGCGCAGCAGT 649

Db 609 TTCAAGAGCAGCAAAATTTATACAACTAATAAAGAGTGATCTCTTAGCAGGAGG 668
Qy 650 TTTCGCCACCGCAGACGCTTTTCCGACGTGGAAATTTACTTTCATATATCATTTGGAGCTGAG 709
Db 669 TTTGATGGCTGACATATATCCCTTCACTGAAGTTTCTTCATGTGCTCAGTGGAAATGAAG 728
Qy 710 TCAAAACCCAGGCGTTTTCATCAGGAGATTGACGATATATTGAAGAGATTCTTTAATGAA 769
Db 729 GGTAAGATTATGAATGCACACCATAGGATGTCATTGTTGAGAAATGTCATCAATGAG 788
Qy 770 CAC-----AAAGCCCAATAGCCTTTTGAAGCGGATAACTTAATG 808
Db 789 CACAAGAAAAATCTTGCAATTTGGAAACTAATGGAGCGTTAGGAGGTGAAGATTTAAT 848
Qy 809 GATGTTCTATTGAATCTTCAAAAAAATGGAACCTTCCAGTGCACGAGTACAAAAACGAAAGC 868
Db 849 GATGTTCTTCAAGACTTATGATGATGAGGCGCTTCAATTTCTCTATCACCAACGACAC 908
Qy 869 ATCAAGACATCCGTTTTCGCAATTTTACTGCGGAGCGGAAACAACTTTCGAAAGCTACA 928
Db 909 ATCAAGCTATAATTTTGAATTTTCTGCTGCCGGACAGAGACTTCTATCGTCAACAAAT 968
Qy 929 GAATGGGTAAATGCGAGAGCTGATGAAATCCAACTGAACAAAGAAAGCACAAGAA 988
Db 969 GTGTGGGCTATGTTAGAAATGTTGAAATCCAGCCGTTATTCGCGAAAGCTCAAGCAGAA 1028
Qy 989 GTTAGACAAGTATTTGGTGAATGGAAGAACTTGAATCAAGATTTTCATGATTTGAAA 1048
Db 1029 GTAGAGAGACTTTAGAGGAAAGAACTTTTCATGAAATGATGTGGAGGAGCTAAAC 1088
Qy 1049 TCTTCAAGTTAGTGGTTAAAGAACTTGAAGATTACATCTCCGTT---GTCTTGAAT 1105
Db 1089 TACCTAAAGTTAGTAATAAAGAACTCTAAGACTTCAAGACTTCCACCGGTTCCACTTTTGCTC 1148
Qy 1106 CCGAGGAGTCTAGAGAAACACAGATTTGATGATGATGAAATTCATCCGAACTCGA 1165
Db 1149 CCAAGAGAAATGTAGGAAAGAGACAAATATAAACCGCTACACTATTCCTGTAAAGACCAA 1208
Qy 1166 ATTGTTGTAATGCTTGGCGGATAGGAAGAGATCTTAATCTTGTGCGAACCTGGAAG 1225
Db 1209 GTCATGTTAATGTTTGGGCTTTGGAGAGATCCAAATATTTGGAATGACGACAGAACT 1268
Qy 1226 TTTAACCCAGAAAGTTTAAAGATTGTCAAATTGATTAATAAGGACGACATTTGAACTG 1285
Db 1269 TTTATGCCAGAGAGATTTGAGCAGTGTCTAAGGATTTTGTGTAATAATTTTGAATAT 1328
Qy 1286 GTACCATTTGTCGAGGAAAGAAATATGCTCTGCAATTTACTTTCAGCTATTACCAATTTG 1345
Db 1329 CTTCATTTGTTGGCGAAGGAGGATTTGCTCTGGGATTTGTTTGGCTTAGCTAATGCT 1388
Qy 1346 GAGTATGTCATTATAAATCTATTATATCAATTTTAAATGGGAACTGGCGGATGGAATACA 1405
Db 1389 TATTTGCCATTTGGCTCAATTAATCACTTTCGATTTGGAACTCCCTGCTGGAATCGAA 1448
Qy 1406 CCTCAACACATTTGATATGATGAGCTATTGGCGGTGCTCTCAGGAAAGAAATAGATCTT 1465
Db 1449 CCAAGCGCATTTGGACTTCACTGAGTTGGTGGAGTAACTGCGCGTAGAAAAAGTGACCTT 1508
Qy 1466 AAGTTGATTTCTATTTCATATCAA 1489
Db 1509 TACTTGGTGGACTCTCTTATCAA 1532

RESULT 9

ADN10626

ID ADN10626 standard; cDNA; 1673 BP.

XX AC ADN10626;

XX DT 15-JUL-2004 (first entry)

XX DE Nicotiana p450 enzyme encoding cDNA SEQ ID NO:211.

XX plant; p450; enzyme; Nicotiana; plant growth regulant; gene therapy;
KW phenotype; gene; ss.
XX Nicotiana sp.
XX WO2004035745-A2.
XX 29-APR-2004.
XX 16-OCT-2003; 2003WO-US032722.
XX 16-OCT-2002; 2002US-0418933P.
PR 08-JUL-2003; 2003US-0485368P.
PR 18-SEP-2003; 2003US-0503989P.
XX (USSM-) US SMOKELESS TOBACCO CO.
XX Xu D;
XX WPI; 2004-348441/32.
DR P-PSDB; ADN10627.
XX New isolated nucleic acid molecules and encoded cytochrome P450 enzymes
PT from Nicotiana plants, useful for altering plant phenotypes.
XX Claim 1; SEQ ID NO 211; 198pp; English.
XX The present sequence represents a plant p450 enzyme nucleic acid molecule
CC isolated from Nicotiana. Also described: (1) an isolated p450 protein
CC from Nicotiana; (2) a transgenic plant comprising the nucleic acid
CC molecule described above; (3) a method of producing the transgenic plant,
CC comprising: (i) operably linking the above nucleic acid molecule with a
CC promoter functional in the plant to create a plant transformational
CC vector; (ii) transforming the plant with the vector; (iii) selecting a
CC plant cell transformed with the transformation vector; and (iv)
CC regenerating a transformation plant containing the above nucleic acid
CC molecule, where the plant is analysed for the presence of the above
CC nucleic acid sequences. The p450 sequences have plant growth regulant
CC activity, and can be used in gene therapy. Compositions and methods from
CC the present invention are useful for altering plant phenotypes.
XX Sequence 1673 BP; 494 A; 309 C; 356 G; 514 T; 0 U; 0 Other;
SQ
Query Match 23.1%; Score 400.4; DB 12; Length 1673;
Best Local Similarity 57.7%; Pred. No. 4.4e-91;
Matches 810; Conservative 0; Mismatchces 561; Indels 33; Gaps 4;
QY 113 CCTCCAGGCCCATGGAAGTTTCCTATCATAGGTAATCTTCCTCATTTTACTCACTTCT 172
DB 135 CCACCAGGTCCATGGAACTACCAATACTAGGAAGTATGCTTCATATGCT-----TGGT 188
QY 173 GATCTAGGCCATGAACGTTTATAGAGCTTGGCTCAAAATTTATGGACCTGTTATGAGTCTT 232
DB 189 GGACTACACACACATGCTCTAGAGATTTAGCCAAAATAATAGGACCACTTATGCACCTT 248
QY 233 CAAATTTGGCCAGTTTTCAGCTGTGTCATTTCTCAGCTGAGAGCCGCAAGAGCTTATG 292
DB 249 CAAATAGGTGAAGTTTTCGCGTGTGGTGTACTTCTCTGATACGCGCAAGAAGTATTA 308
QY 293 AAAAATCTCAGGCTGATGCCCTTCGCCCAACGCCCTATGCTCTGGAGCGCACAGATTTGTTT 352
DB 309 AAAATCATGACATCGCTTTTCGGCTAGGCTTAGCTTTTGGCCCGGAGATTTGCTGT 368
QY 353 TATAATCGGAAAGATGTCTTTGCTTCATATGAGATCATCTGGAGCGAGATGAAGAAA 412
DB 369 TACAATAGTCTGATCTAGCTAGCTTTTGGCCCTATGGCGACTATTGGAGACAAATGCGTAAA 428
QY 413 ATTTGGATACCTGAATTTCTGATGCCAAAAGTTCAATCTCCAGGTTATCCGAG 472
DB 429 ATATGTGCTTGGAGTGCTAGTGCCAAAGATGTTCCGACATGTTAGTCTTATTAGCGG 488

QY 473 GAAGAAATGGAGATGCCATCATTCTCCGTTGGAAGCCGATCTCCGGTCAATATT 532
DB 489 ATGAAGTTCTTCGCTCATTAATTTTATCCGGTATCTTCTGGTGAACCTATTAATGTT 548
QY 533 ACAAGATCATTTATGGCAATTATAATTTTCGATCATGATAAGAACATCCGTTGGTAA--T 589
DB 549 ACGGAAAGGATCTTTTGTTCACAGCTCCATGACATGTAGATCAGCGTTTGGGCAAGTG 608
QY 590 TGTNAGCAAAAGAAAGATTGCTGAGTGTGCGGATGATCAATGAGGAGCGAGAGT 649
DB 609 TTCAAGAGCAAGACAAATTTATACAACTAATTAAGAAAGTGATATCTTTAGCAGGAGG 668
QY 650 TTTGGCACCGCAGAGCGCTTTTCGACGTCGAAATTAATTTCACTATATCAATTTGAGCTGAG 709
DB 669 TTTGATGGCTGACATATTTCCCTTCACTGAAGTTTCTTCATGTGCTCAGTGGAAAGAAG 728
QY 710 TCAAAACCCAGCGCTTTGATCAGGAGATTTGACGATATATCTTGAAGAGATTTCTTAATGAA 769
DB 729 GGTAAAGATTATGAATGCACACCATAAAGGTAGATGTCATTTGTTGAGAAATGTCATCAATGAG 788
QY 770 CAC-----AAAGCCAAATAGCCCTTTTGAAGCGGATTAATG 808
DB 789 CACAAGAAAAATCTTTGCAATTTGGGAAACTAATGGAGCGTTAGGAGGTGAAGATTTAAAT 848
QY 809 GATGTTCTATTGAATCTTCAAAAAAATGAAAACGTTCCAGTGCACGTGACAAACGAAAGC 868
DB 849 GATGTTCTTAAAGACTTATGAATGATGGAGGCTTCAATTTCTTATACCAACGACAC 908
QY 869 ATCAAGACATCGTTTTCGAAATGTTTACTGCGGAGCGAAACAACTTCGAAAGCTTACA 928
DB 909 ATCAAGAGCCATAATTTTGGACATGTTTGTGCGGAGACAGACTTTCATCGTCAACAAAT 968
QY 929 GAATCGGTAAATGGCAGAGCTGATGAATTAATCAACTGAATGAAGAAAGACAAAGAA 988
DB 969 GTGCGGCTATGGTAGAATTTGTAATAATCCAGACGCTATTCGCGAAAGCTCAAGCAGAA 1028
QY 989 GTTAGACAAGTATTTGGTGAATGGAAATGGAAATGATGAATCAAGATTTTCATGATTTGAAA 1048
DB 1029 GTAAGAGAGCATTTTAGAGAAAGAAACTTTTCATGAATAATGATGTGGAGGAGCTTAAC 1088
QY 1049 TTTCTCAAGTTAGTGGTTTAAAGAACTCTAAGATTTACATCTCCCGTT---GTCTTGATT 1105
DB 1089 TACCTAAAGTTAGTAATAAAGAAACTCTAAGACTTTCATCCACCGGTTCCACTTTTGCTC 1148
QY 1106 CCGAGGAGCTGAGAGAAACACAGAAATGATGATGATGAATTAATTCATCCGACACTCGA 1165
DB 1149 CCAAGAGAAATGAGGAAAGAGACAAATATAAACGGCTACACTATTCTCTGTAAGACCAA 1208
QY 1166 ATTTGTTGTAATGCTTTGGCGGATAGGAAGAGATCCCTAATCTTTGTCGGAACCTGGAAG 1225
DB 1209 GTCATGTTAATGTTTGGCTTTGGAGAGATCCAAATATTTGGAATGACGAGAACT 1268
QY 1226 TTTAAACCAGAAAGGTTTAAAGATTGTGCAATTTGATTAATAAGGAGCAGCATTTGAACGTG 1285
DB 1269 TTTATGCCAGAGATTTGAGCAGTGTCTTAAGGATTTTGTGGTAATAATTTTGAATAT 1328
QY 1286 GTACCATTGTTGTCAGGAAAGAGATATGCTCTGCGCATTTACTTTCAGCTATTACCAATTTG 1345
DB 1329 CTTCCATTGTTGGCGGAGGAGGATTTGCTCTGGGATTTCTGTTGGCTTAGCTAATGCT 1388
QY 1346 GAGTATGCTATTATAAATCTATTATATCAATTTTAAATGGAACTGGCCGATGAATTTACA 1405
DB 1389 TATTTGCCATTGGCTCAATTACTATATCACTTCGATTTGGAATCTCCCTGCTGGAACTGAA 1448
QY 1406 CCTCAAAACATTTGATATGACTGAAGCTATTTGGCGGTGCTCTCAGGAAAAAATAATGATCTT 1465
DB 1449 CCAAGCGACTTGGACTTGAATGAGTGGTGGAGTAACTGCCGCTAGAAAAAGTGACCTT 1508
QY 1466 AAGTTGATTTCTTATCCATATCAA 1489
DB 1509 TACTTGGTTGGACTCTCTTATCAA 1532

```
RESULT 10
ID ADZ65334 standard; cDNA; 1673 BP.
XX
XX
AC ADZ65334;
XX
XX 14-JUL-2005 (first entry)
XX
XX Tobacco full length cDNA for cytochrome p450 clone D209-AH12.
XX
XX Cytochrome p450; ss; secondary metabolite; ethylene; plant; senescence;
KW nornicotine; transgenic plant; gene.
XX
XX Nicotiana tabacum.
XX
XX WO2005038033-A2.
XX
XX 28-APR-2005.
XX
XX 15-OCT-2004; 2004WO-US034065.
XX
XX 16-OCT-2003; 2003US-00686947.
XX
XX 29-APR-2004; 2004US-0568235P.
XX
XX 03-SEP-2004; 2004US-00934944.
XX
XX (USM-) US SMOKELESS TOBACCO CO.
XX
XX Xu D;
XX
XX WPI; 2005-315717/32.
XX
XX P-PSDB; ADZ65335.
XX
XX New nucleic acid molecule encoding cytochrome p450 enzymes in Nicotiana,
XX useful in developing tobacco plants with altered phenotypes.
XX
XX Disclosure; SEQ ID NO 231; 226pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) from
XX Nicotiana, where the nucleic acid molecule comprising any of the 59
XX CC nucleic acid sequences of SEQ ID NOS: 299-357 (NOTE: The claims refer to
XX CC SEQ ID NOS 299-357 as nucleic acids but these sequences (apart from SEQ
XX CC ID NO 356) are all proteins and appear as ADZ65402-ADZ65460. The nucleic
XX CC acids of the invention encode cytochrome p450 enzymes whose expression is
XX CC induced by ethylene and/or plant senescence. Also included are a
XX CC transgenic plant comprising (I), a method of producing a transgenic
XX CC plant, a method of selecting a plant containing a nucleic acid molecule
XX CC (where the plant is analyzed for the presence of nucleic acid sequence of
XX CC ADZ65402-
XX
XX Sequence 1673 BP; 494 A; 309 C; 356 G; 514 T; 0 U; 0 Other;
XX
XX Query Match 23.1%; Score 400.4; DB 14; Length 1673;
XX Best Local Similarity 57.7%; Pred. No. 4.4e-91;
XX Matches 810; Conservative 0; Mismatches 561; Indels 33; Gaps 4;
XX
XX QY 113 CCTCAGGGCCATGGAAGTTTCTATCATAGTAATCTTCTCATTTATCTACTCTCT 172
XX Db 135 CCACCAGGTCATGGAACTACCAATACCTAGGAAGTATGCTTCATATGTT-----TGGT 188
XX
XX QY 173 GATCTAGGCCATGAACGTTTATAGAGCTTGGCTCAAAATTTATGACCTGTTATGAGTCTT 232
XX Db 189 GGACTACCAACCATGTCCTTAGAGATTTAGCCAAAATATATGGACCACTTATGCACCTT 248
XX
XX QY 233 CAAATTGGCCAAAGTTTTCAGCTGTGTGTCATTTCTTCTAGCTGAAGCAGCCAAAGAGTTATG 292
XX Db 249 CAATTAGGTGAAGTTTCTCGGTTGGTTGTTACTTCTCTGTATACGGCAAAAGATTA 308
XX
XX QY 293 AAAAATCAGGCTGATGCCCTTCGCCCAACGCCCTATCGTCTTGGACGCAAGATTTGTTT 352
XX Db 309 AAAAATCATGATCGCTTTTGGCTTAGGCCCTAGCCCTTTTGGCCCGGAGATTTGTCGT 368
XX
XX QY 353 TATATCGGAAGATGCTTGTGTTGCTTCATATGGAGATCACTGGAGGAGATGAAGAA 412
XX Db 113 TACAATAGGCTGATCTAGCCCTTTTGCCCTATGGGACTATTTGGAGACAATGCGTAA 428
XX QY 413 ATTTGGATACCTTGAATTTCTGAGTGCCAAAGAAAGTTCAATCCTCAGGTTAATCCGAGAG 472
XX Db 429 ATATGTGCTTGGGAAGTGCTCAGTGCCAAAGAAAGTTTGGACATTTAGCTCTATTAGGCGG 488
XX
XX QY 473 GAAGAAATGGAGGATGCCATCACATTCCTCGTTTGGAAAGCCGATCCTCCGGTCAATATT 532
XX Db 489 AATGAAGTTCTTCTGCTCATTAATTTTATCCGGTCATCTCTCTGGTGAACCTATTTAATGTT 548
XX
XX QY 533 ACAAAGATCATTTTATGGCATTATTAATTTCCGATCATGATAAGAAATCCGTTGGTAA---T 589
XX Db 549 ACGAAAGGATCTTTTGTTCACAGCTCCATGACATGATAGATCAGCGTTTGGGCAAGTG 608
XX
XX QY 590 TGTAAAGCAAAAGAAAGATTGCTGAGTGTTCGCCATGCGATCAATGAGGCGACGACAGT 649
XX Db 609 TTCAAAGAGCAAGACAAATTTATACAACTAATTAAGAAAGTGATACTCTTAGCAGAGGG 668
XX
XX QY 650 TTTGGCACCGCAGAGCGTTTTCGACGTGGGAAATTTACTTCACTATATCATTTGGAGCTGAG 709
XX Db 669 TTTGATGTGGCTGACATATTCCTTCACTGAAGTTTCTTCATGTGCTCAGTGGAAATGAG 728
XX
XX QY 710 TCAAAACCCAGCGGTTTTCATCAGGAGATTGACGATATACCTTGAAGAGATTTCTTAATGAA 769
XX Db 729 GGTAAAGATTGAATGCACACCATTAAGTAGATGCCATTGTTGAGAATGTCATCAATGAG 788
XX
XX QY 770 CAC-----AAAGCCAAATAGCCCTTTTGAAGCGGATAAATCTTTAATG 808
XX Db 789 CACAAGAAAAAATCTTGGCAATTTGGGAAACTAATGGAGCGTTTAGGAGGTGAAGATTTAAT 848
XX
XX QY 809 GATGTTCTATTGAATCTTCAAAAAAATGGAAACGTTCCAGTGCCAGTCACAAACCAAGAC 868
XX Db 849 GATGTTCTTCAAGCATTTATGAATGATGGAGCCCTTCAATTTCTTATCACCACCAACGAC 908
XX
XX QY 869 ATCAAAGCATCCGTTTTCGAAATGTTTACTGCGGGAGCGAAACAACTTCGAAAGCTACA 928
XX Db 909 ATCAAAGCCATAATTTTGGACATGTTTCTGCGGGACAGAGACTTCATCGTCAACAAAT 968
XX
XX QY 929 GAATGGGTAATGGCAGAGCTGATGAAATCCAACTGAACCTAAGAAAAGCACAAGAA 988
XX Db 969 GTGTGGGCTATGTTAGAAATGGTGAATAATCCAGCCGTTATTCGCGAAAGCTCAAGCAGAA 1028
XX
XX QY 989 GTTACAGCAAGTATTTGGTGAATGGAAAGTTGATGATCAAGATTTTCATGATTGAAA 1048
XX Db 1029 GTAAAGAGAGCATTTTAGAGGAAAGAAACTTTTCATGAAATATGATGCGAGGAGCTAAAC 1088
XX
XX QY 1049 TTCTTCAAGTTAGTGGTTTAAAGAAACTCTAAGATTACATCTCTCCGGTT---GTCTGATT 1105
XX Db 1089 TACCTAAGTTAGTAAATAAAGAAACTCTAAGACTTCATCCACCGGTTCCACTTTTGCTC 1148
XX
XX QY 1106 CCGAGGGAGTGTAGAGAAACAAACAGAAATTTGATGATATGAATTTATCCGACACTCGA 1165
XX Db 1149 CCAAGAGAATGTAGGAAAGAGACAAATATAAACCGCTACACTATTCTGTAAAGAACCAAA 1208
XX
XX QY 1166 ATTGTTGTAATGCTTGGGCGATAGGAAGAGATCCTAATCTATCTTGGTCGNACTCGAAAG 1225
XX Db 1209 GTCATGTTAATGTTTGGGCTTTGGGAAGAGATCCAAATATTTGGAATGACCGCAAACT 1268
XX
XX QY 1226 TTTTAAACCAGAAAGGTTTAAAGATTGTGCAATTTGATTAATAAGGGACGACATTTGAAC 1285
XX Db 1269 TTTATGCCAGAGAGATTTGACGAGTGCTCTAAGGATTTTGTGGTAAATAATTTGAATAT 1328
XX
XX QY 1286 GTACATTTGGTGAGGAAAGAAATATGTCCTGGCATTACTTCACTATTTACCAATTTG 1345
XX Db 1329 CTTCCATTGTTGGTGGGAGGAGGAGGATTTGTCCTGGGATTTCTGTTGGCTTAGCTAATG 1388
XX
XX QY 1346 GAGTATGTCATTATAAATCTATTATATCATTTTAAATTTGGAACTGGCCGATGGAATTACA 1405
XX Db 1389 TATTTGGCAATGGCTCAATTAATCTATATCACTTCGATTTGGAACTCCCTGCTGGAATCG 1448
XX
XX QY 1406 CCTCAAAACATTTGATATGACTGAAGCTATTTGGCGGTGCTCTCAGGAAAAAATATGATCT 1465
XX Db 1449 CCAACGCACTTGGACTTGAATTTGGTGGAGTAACTGCGCGCTAGAAAAAAGTGACCTT 1508
```

QY 1466 AAGTTGATTCCTATTCCATACAA 1489
 |||||
 Db 1509 TACTTGGTTCGACTCCTTATCAA 1532

RESULT 11

ADZ65312
 ID ADZ65312 standard; cDNA; 1673 BP.

XX
 AC ADZ65312;
 XX

DT 14-JUL-2005 (first entry)

XX Tobacco full length cDNA for cytochrome p450 clone D209-AA10.

XX Cytochrome p450; ss; secondary metabolite; ethylene; plant; senescence;
 KW nornicotine; transgenic plant; gene.

XX Nicotiana tabacum.

XX WO2005038033-A2.

XX 28-APR-2005.

XX 15-OCT-2004; 2004WO-US034065.

XX 16-OCT-2003; 2003US-00686947.

PR 29-APR-2004; 2004US-0566235P.

PR 03-SEP-2004; 2004US-00934944.

XX (USSW-) US SMOKELESS TOBACCO CO.

XX Xu D;

XX WPI; 2005-315717/32.

DR P-PSDB; ADZ65313.

XX New nucleic acid molecule encoding cytochrome P450 enzymes in Nicotiana,
 PT useful in developing tobacco plants with altered phenotypes.

XX Disclosure; SEQ ID NO 209; 226pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) from
 CC Nicotiana, where the nucleic acid molecule comprising any of the 59
 CC nucleic acid sequences of SEQ ID NOS: 299-357 (NOTE: The claims refer to
 CC SEQ ID NOS 299-357 as nucleic acids but these sequences (apart from SEQ
 CC ID NO 356) are all proteins and appear as ADZ65402-ADZ65460. The nucleic
 CC acids of the invention encode cytochrome p450 enzymes whose expression is
 CC induced by ethylene and/or plant senescence. Also included are a
 CC transgenic plant comprising (I), a method of producing a transgenic
 CC plant, a method of selecting a plant containing a nucleic acid molecule
 CC (where the plant is analyzed for the presence of nucleic acid sequence of
 CC ADZ65402-
 XX

SQ Sequence 1673 BP; 493 A; 309 C; 357 G; 514 T; 0 U; 0 Other;

Query Match 23.1%; Score 400.4; DB 14; Length 1673;

Best Local Similarity 57.7%; Pred. No. 4.4e-91;

Matches 810; Conservative 0; Mismatches 561; Indels 33; Gaps 4;

QY 113 CCTCAGGGCCATGGAAGTTTCCTATCATAGTAAATCTTCCTCATTTTACTCACTTCT 172

|||
 Db 135 CCACAGGTCCATGGNACTACCAATACTAGGAAGTATGCTTCATATGTT-----TGGT 188

QY 173 GATCTAGGCCATGAACGTTTTAGAGCCTTGGCTCAAAATTTATGGACCTGTTATGAGTCTT 232

|||
 Db 189 GCACTACCAACACCATGCTCCTTAGAGATTTAGCCAAAATAATGGACCCTTAATGACCTT 248

QY 233 CAAATGGCCAGATTTCAGCTGTTGTCATTTCTTCAGCTGAAGCAGCCAAAGAGTTATG 292

|||
 Db 249 CAAATAGGTGAAGTTTCTCGGTTGTTGTTACTTCTCTCATGATACGCGCAAAAGATTTA 308

QY 293 AAAAATCAGGCTGATGCTTTCGCCAAACGCCCTTATCGTCTTGGAGCACACAGATTGTGTTT 352
 |||||
 Db 309 AAAATCTCATGACATCGCTTTTGGCTTAGGCTTAGCTTTTGGCCCCGGAGATTGCTGT 368
 |||||
 QY 353 TATAATCGGAAAGATGTCCTTGTGTTTGGTTTCAATGAGAGATCACTGGAGGAGATGAAGAAA 412
 |||||
 Db 369 TACAATAGGTCTGATCTAGCTTTTGGCCCTATGCGGACTATTGGAGACAAAATGCGTAAA 428
 |||||
 QY 413 ATTGGATACCTTGAATTTCTGAGTGCCAAAAGTTTCAATCTCTCCAGGTTAATCCGAGAG 472
 |||||
 Db 429 ATATGTGCTTGGAGTGTCTCAGTGCCAAAGATGTTTCGGACATTTAGTCTCTATTAGCGG 488
 |||||
 QY 473 GAAGAAATGGAGATGCCATCACATTCCTCCGTTCCGAAAGCCGGATCTCCGGTCAATATT 532
 |||||
 Db 489 AATGAAGTTCTTCGCTCTCATTAATTTATCCGGTCACTTCTGTTGTAACCTATTAATGTT 548
 |||||
 QY 533 ACAAGATCATTTATGGCATTTAATTTTCGATCATGATAAGAACATCCGTTGTGTA--T 589
 |||||
 Db 549 ACGGAAAGGATCTTTTGTTCACAGCTCCATGACATGTAGATCAGCGTTTGGGCAAGTG 608
 |||||
 QY 590 TGTAAAGCAAAAGAAAGATTGCTGAGTGTGCGGATGCAGTCAATGAGGCGACGAGT 649
 |||||
 Db 609 TTCAAAGAGCAAGACAAATTTATACAACTAAATTAAGAAAGTGATACTCTTAGCAGGGG 668
 |||||
 QY 650 TTTGGCACCGCAGACGCTTTTCGAGCTGGAATTTACTTCACTATATCATTTGAGGCTGAG 709
 |||||
 Db 669 TTTGATGTGGCTGACATATTCCTTCACTGAAGTTTCTTCATGTGCTCAGTGAATGAAG 728
 |||||
 QY 710 TCAAAACCCGAGCGTTTGCATCAGGAGATTGACGATATATCTTGAAGAGATTCTTAATGAA 769
 |||||
 Db 729 GGTAAAGATTATGAATGCAACCATTAAGGTAGATGCCATTTGTTGAGAAATGTCATCAATGAG 788
 |||||
 QY 770 CAC-----AAAGCCAAATAGCCCTTTTGAAGCGGATAACTTAATG 808
 |||||
 Db 789 CACAAGAAAAATCTTGCAATTTGGGAAAACTAAATGGAGCGTTAGGAGGTGAAGATTTAAT 848
 |||||
 QY 809 GATGTTCTTATGCAATCTTCAAAAAAATGAAAAAGTTCCAGTGCACGTGACAAACGAAAGC 868
 |||||
 Db 849 GATGTTCTTAAAGACTTATGAATGATGGAGGCTTCAATTTCTCTATCACCACGACAAAC 908
 |||||
 QY 869 ATCAAAGCATCCGTTTGGCAATTTTACTGCGGAGGCGAAACAACTTCGAAAGGTACA 928
 |||||
 Db 909 ATCAAAGCTATAAATTTTGGACATGTTTGTGCGGAGCAGAGACTTCTATCGTCAACAAT 968
 |||||
 QY 929 GAATCGGTAAATGGCAGAGCTGATGAATTAATCAACTGAATGAAGAAACACAGAGAA 988
 |||||
 Db 969 GTGTCGGCTATGGTGAATTTGGTGAATTTCCAGCGGTATTCGGAAGGCTCAAGCAGAA 1028
 |||||
 QY 989 GTTAGACAAGTATTTGGTGAATTTGGAAATGGAATGATGAATCAAGATTTTCATGATTTGAAA 1048
 |||||
 Db 1029 GTAAGAGAGCATTTTAGAGGAAAGAAACTTTCGATGAATGAATGATGGAGGAGCTTAAC 1088
 |||||
 QY 1049 TTCTTCAAGTTAGTGGTTTAAAGAAACTCTAAGATTACATCTCCGTT---GTCTTGAATT 1105
 |||||
 Db 1089 TACCTAAAGTTAGTAATAAAGAAACTCTAAGACTTCTCCACCGGTTCCACTTTTGTCTC 1148
 |||||
 QY 1106 CCGAGGAGGTAGAGAAACACACGAATTTGATGATGAATTCATCCGACACTCGA 1165
 |||||
 Db 1149 CCAAGAGAAATGTAGGGAAGAGACAAATATAAACCGCTACACTATTCTCTGTAAGACCAA 1208
 |||||
 QY 1166 ATTGTGTGAATGCTTTGGCGGATAGGAAGATCTTAATCTTGGTCGGAACCTGGAAG 1225
 |||||
 Db 1209 GTCATGTTAATGTTTGGCTTTGGAGAGAGATCCAAATATTTGAATGACGAGAACT 1268
 |||||
 QY 1226 TTTAACCCAGAAAGTTTAAAGATTGTGCAATTTGATTAATAAGGAGCAGCATTTTGAATG 1285
 |||||
 Db 1269 TTTATGCCAGAGAGATTTGAGCAGTGTCTAAGAGATTTTGTGGTAAATAATTTTGAATAT 1328
 |||||
 QY 1286 GTACCATTTGGTCAGGAAAGAGATATGTCCTGCGCATTTACTTCAGCTATTTACCATTTG 1345
 |||||
 Db 1329 CTTCCATTTGGTGGCGGAAGGAGATTTGTCTCGGATTTCTGTTGGCTTAGCTAATGCT 1388
 |||||
 QY 1346 GAGTATGTCATTATAAATCTATTATATCATTTTAAATTTGGAACTGCGCCGATGGAATTA 1405
 |||||

Db 1389 TATTTCGCAATGGCTCAATTAATCACTTCAATTCGATTGGAACTCCCTGCTGGAAATCGAA 1448
Qy 1406 CCTCAACACATTGATATCACTGAAGCTATTGGCGGTGCTCTCAGGAAAAAATAGATCTT 1465
Db 1449 CCAAGCGACTTGGACTTGACTGAGTTGGTGGAGTAACCTGCGCTAGAAAAAAGTGACCTT 1508
Qy 1466 AAGTTGATTTCCTATTCCATATCAA 1489
Db 1509 TACTTGGTTGCGACTCCTTATCAA 1532

RESULT 12

AD265314
ID AD265314 standard; cDNA; 1673 BP.
XX AC AD265314;
XX DT 14-JUL-2005 (first entry)
XX DE Tobacco full length cDNA for cytochrome p450 clone D209-AA12.

XX KW Cytochrome p450; ss; secondary metabolite; ethylene; plant; senescence;
XX KW normicotine; transgenic plant; gene.

XX OS Nicotiana tabacum.

XX PN WO2005038033-A2.

XX XX 28-APR-2005.

XX 15-OCT-2004; 2004WO-US034065.

XX 16-OCT-2003; 2003US-00686947.

XX 29-APR-2004; 2004US-0566235P.

XX 03-SEP-2004; 2004US-00934944.

XX (USSM-) US SMOKELESS TOBACCO CO.

XX Xu D;

XX WPI; 2005-315717/32.

XX P-PSDB; AD265315.

XX New nucleic acid molecule encoding cytochrome p450 enzymes in Nicotiana,
XX useful in developing tobacco plants with altered phenotypes.

XX Disclosure; SEQ ID NO 211; 226pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) from
XX Nicotiana, where the nucleic acid molecule comprising any of the 59
XX nucleic acid sequences of SEQ ID NOS: 299-357 (NOTE: The claims refer to
XX SEQ ID NOS 299-357 as nucleic acids but these sequences (apart from SEQ
XX ID NO 356) are all proteins and appear as AD265402-AD265460. The nucleic
XX acids of the invention encode cytochrome p450 enzymes whose expression is
XX induced by ethylene and/or plant senescence. Also included are a
XX transgenic plant comprising (I), a method of producing a transgenic
XX plant, a method of selecting a plant containing a nucleic acid molecule
XX (where the plant is analyzed for the presence of nucleic acid sequence of
XX AD265402-
XX

XX SQ Sequence 1673 BP; 494 A; 309 C; 356 G; 514 T; 0 U; 0 Other;

Query Match 23.1%; Score 400.4; DB 14; Length 1673;

Best Local Similarity 57.7%; Pred. No. 4.4e-91;

Matches 810; Conservative 0; Mismatches 561; Indels 33; Gaps 4;

Qy 113 CCTCAGGCGCCATGGAAGTTTCCTATCATAGGTAATCTTCCTCATATTACTCACTTCT 172

Db 135 CCACCAGGTCCTATGGAACCTACCACTACTAGGAAGTATGCTTCATATGCT-----TGGT 188

Qy 173 GATCTAGGCCATGAACGCTTTTAGAGCTTGGCTCAAAATTATGGACCTGTTATGAGTCTT 232

Db 189 GGACTACCAACCACATGCTCTTAGAGATTAGCCAAATAATATGACCACCTTATGCACCTT 248
Qy 233 CAAATTGGCCAAAGTTTCAGCTGTTGTCAATTTCTTCAGCTGAAGCAGCCAAAGAGGTTATG 292
Db 249 CAA'TAGGTGAAGTTTCTGCGGTTGTGGTTACTTCTCTGTAGTCGCAAAAGAAATATTA 308
Qy 293 AAAACTCAGGCTGATGCTTTGCGCCAAAGCCTATCGTCTTGGAGCGCACAGATCTGTGTTT 352
Db 309 AAAACTCATGACATCGCTTTTGGCTTAGGCTAGCCCTTTTGGCCCCGGAGATTGCTGT 368
Qy 353 TATAATCGMAAGATGCTCTTCTTGTCTTATATGGAGATCACTGGAGCGAGATGAAGAAA 412
Db 369 TACAATAGGCTGATCTAGCCCTTTTGGCCCTATGGCAGCTATGGAGACAAATGCGTAAA 428
Qy 413 ATTTGGATACATTGTTGAGTGCCAAAGTTTCAATCTCAGGTTTAAATCCGAGAG 472
Db 429 ATATGTGCTTGGAAAGTGCTCAGTGCCAAAGAAATGTTTCGGACATTTAGCTCTATTAGGCGG 488
Qy 473 GAAGAAATGGAGGATGCCATCACATTCCTCGTTTCGAAAGCGGATCTCCGGTCAATATT 532
Db 489 AATGAAGTTCTTCGTTCTCAITTAATTTTATCCGGTCACTCTCTGGTGAACCTATTAAATGTT 548
Qy 533 ACAAAGATCAATTTATGGCAATTAATTTTCGATCATGATAAGAACAATCCGTTGGTAA---T 589
Db 549 ACGGAAGGATCTTTTGTTCACAAGCTCCATGACATGTAGATCAGCGTTTGGGCAAGTG 608
Qy 590 TGTAAAGCAAAAGAAAGATTGCTGAGTGTGCCGATGCGAGTCAATGAGGCGAGCGAGT 649
Db 609 TTCAAGAGCAAGACAAATTTATACAACTAATAAAGAAAGTGATCTCTTAGCAGGAGGG 668
Qy 650 TTTGGCCCGCAGAGCGCTTTCCGACGTGGAAATTTACTTTCATATATCATTTGGAGCTGAG 709
Db 669 TTTGATGGCTGACATATTCCTTCACTGAAGTTTCTTCATGTCTCAGTGAATGATGAG 728
Qy 710 TCAAAACCCAGCGCTTTGTCATCAGGAGATTGACGATATATCTTTGAAGAGATTCTTTAATGAA 769
Db 729 GGTAAAGATTATGAATGCACACCAATAAGGTAGTCCATTTGTTGAGAAATGTCATCAATGAG 788
Qy 770 CAC-----AAAGCCAAATAGCCTTTTGAAGCGGATCACTTAATG 808
Db 789 CACAAGAAAAATCTTGCAAATTTGGGAAAACTAATGGAGCGTTAGGAGGTGAAGATTAAAT 848
Qy 809 GATGTTCTATTGAATCTTCAAAAAAATGMAAACGTTCCAGTGCCAGTGACAAACGAAGAC 868
Db 849 GATGTTCTTAAAGACTTATGAATGATGAGGCGCTTCAATTTCTCATGTCTCAGTGAATGAG 908
Qy 869 ATCAAAGCATCCGTTTTCGCAAAATGTTTACTGCGGGAGCGAAACAACTTCGAAAGCTACA 928
Db 909 ATCAAAGCCATAATTTTGGACATGTTTCTGCGGGACAGAGACTTCATCGTCAACAAAT 968
Qy 929 GAATGGGTAATGGCAGAGCTGATGAAAAATCCAATCACTGAATGAAGAAAAAGCACAAGAGAA 988
Db 969 GTGTGGGCTATGTTAGAAATGTTGAAAAATCCAGCCGTTATTCGCAAAAGCTCAAGCAGAA 1028
Qy 989 GTTAGACAAGTATTTGGTGAATGGGAAAGTTGATGAATCAAGATTTCATGATTGAAA 1048
Db 1029 GTAAGAGAAGCACTTTAGAGGAAAAAGAACTTTTCAATGAAAAATGATGTGGAGAGCTAAAC 1088
Qy 1049 TTTCTCAAGTTAGTGGTTAAAGAAACTCTTAAGATTACATCTCCCGGTT---GTCTTGATT 1105
Db 1089 TACCTAAAGTTAGTAAATAAAGAAACTCTAAGACTTCATCCACCGGTTCCACTTTTGCCTC 1148
Qy 1106 CCGAGGAGGTGATAGAGAAACAAACGAATTTGATGATATGAATTCATCCGACACTCGA 1165
Db 1149 CCAAGAGAATGTAGGGAAGAGACAAATATAAACCGCTACACTATTCTCTGTAAGACCAAA 1208
Qy 1166 ATTTGTTGTAATGCTTTGGGCGATAGAGAGATCTTAATACATTGTTGGTGGAACTCGTGAAG 1225
Db 1209 GTCTGGTTAAATGTTTGGGCTTTGGGAAGAGATCCAAATATTTGAAATGACGAGAACT 1268
Qy 1226 TTTAAACCCAGAAAGGTTTAAAGATTGTGCAATTTGATTTAAAGGGACGACATTTGAACGTG 1285
Db 1269 TTTATGCCAGAGAGATTGAGCAGTGCCTAAGGATTTTGGTGGTAAATAATTTTGAATAT 1328

Db 789 CACAAGAAAAATCTTGCATTTGGAAAACTAATGGAGCGTTAGGAGTGAAGATTAAATT 848
QY 809 GATGTTCTTATGTAATCTTCAAAAAATGGAAAGTTCCAGTGCAGTGACAAACGAAGC 868
Db 849 GATGTTCTTCTAAGACTTATGAATGATGGAGGCTTCAATTTTCTATCAACACGACAAC 908
QY 869 ATCAAGACATCCGTTTTCGAAATGTTTACTGCGGGAGCGAACAACCTTCGAAAGCTACA 928
Db 909 ATCAAGACTATAATTTTGACATGTTTGTGCGGGACAGAGACTTCATGTCACAAATT 968
QY 929 GAATGGGTATGCGAGCTGATGAAAAATCCAACTGAACCTAAGAAAGCACAAGAGAA 988
Db 969 GTGTGGCTATGCTAGAAAATGGTGAATAATCCAGCCGTATTTCGCGAAAGCTCAAGCAGAA 1028
QY 989 GTTAGACAAGTATTTGGTGAATGGGAATGGGAAGTTGATGAATCAAGATTTCATGATTTGAAA 1048
Db 1029 GTAAGAGAGCAATTTAGAGGAAAAGAACTTTTCGATGNAATGATGCGAGGAGCTAAAC 1088
QY 1049 TTCTTCAAGTTAGTGTGTTAAAGAACTCTAAGATTACATCTCCGGTT--GCTTGAAT 1105
Db 1089 TACCTAAGTTAGTATAAAGAAACTCTAAGACTTCATCCACCGTTCCACTTTTGCTC 1148
QY 1106 CCGAGGAGTGTAGAAAAACAACAGAAATGATGGATATGAAATTCATCCGAACCTCGA 1165
Db 1149 CCAAGAGAAATGTAGGAAGAGACAAATATAAACCGCTACACTATTCTGTAAAGACCAA 1208
QY 1166 ATGTTGTGAATGCTTGGCGGATAGGAAGATCTTAATCTTGGTGGGAACCTGGAAG 1225
Db 1209 GTCATGGTAAATGTTTGGGCTTTGGGAAGAGATCCAAATAATTTGGAATGACGAGAACT 1268
QY 1226 TTTAACCCAGAAAGTTTAAAGATTGTGCAATTGATTATAAGGGACGACATTGAACTG 1285
Db 1269 TTATGCCAGAGAGATTGCGAGTCTCTAAGGATTTGTTGGTAAATTTGATAT 1328
QY 1286 GTACATTTGGTGCAGAAAAAGAAATGATGCTCGGCATTTACTTCACTAATACCAATTTG 1345
Db 1329 CTTCCATTTGGTGGCGAAGGAGGATTTGCTCGGATTTCTGCTGCTAGCTAATGCT 1388
QY 1346 GAGTATGTCATTATAATCTATTATATCATTTTAAATGGGAATGCGCGATGGAATTACA 1405
Db 1389 TATTTGCCAATGGCTCAATTTACTATATCACTTCGATTGGAATCTCCCTGCTGGAATCGAA 1448
QY 1406 CTTCAAAACATTTGATGATGACTGAAGCTATTGGCGGTGCTCTCAGGAAAAAATAGATCTT 1465
Db 1449 CCAAGCACTTGGACTTGACTGATGTTGGTGGATGATGCTGCGCTAGAAAAAGTGACCTT 1508
QY 1466 AAGTTGATTCCTATTCATATCAAA 1489
Db 1509 TACTTGGTTGCGACTCCTTATCAA 1532

RESULT 15

ADZ65881
ID ADZ65881 standard; cdNA; 1673 BP.
XX
AC ADZ65881;
XX

DT 14-JUL-2005 (first entry)

DE Tobacco cytochrome P450 enzyme cdNA #116.

KW Enzyme engineering; cytochrome P450; gene; ss.

XX Nicotiana tabacum.

XX WO2005038018-A2.

XX 28-APR-2005.

PF 15-OCT-2004; 2004WO-US034218.

XX 16-OCT-2003; 2003US-00686947.

PR 29-APR-2004; 2004US-0566235P.
PR 17-SEP-2004; 2004US-00943507.
XX (USM-) US SMOKELESS TOBACCO CO.
XX Xu D;
XX
DR WPI: 2005-315709/32.
DR P-PSDB; ADZ65882.
XX
PT New isolated nucleic acid molecule from Nicotiana, useful for altering
PT plant phenotypes, thus producing a transgenic plant having reduced levels
PT of nicotine.
XX
PS Disclosure; SEQ ID NO 231; 203pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule from
CC Nicotiana, encoding a protein. The invention also relates to a transgenic
CC plant comprising the nucleic acid molecule, a method of producing a
CC transgenic plant comprising operably linking the nucleic acid molecule
CC with a promoter functional in the plant to create a plant
CC transformational vector, transforming the plant with the plant
CC transformational vector, selecting a plant cell transformed with the
CC transformational vector and regenerating a transformation plant from the
CC transformed plant cell, a method of selecting a plant containing a
CC nucleic acid molecule, a method of increasing or decreasing nicotine
CC levels in a plant by operably linking the nucleic acid molecule with a
CC promoter functional in the plant, a tobacco product having reduced
CC amounts of nicotine levels, the tobacco product comprising tobacco
CC from the plant, a tobacco leaf having reduced amounts of nicotine
CC levels and a method of isolating a gene from a plant using the isolated
CC nucleic acid. In producing a transgenic plant, the plant has reduced
CC levels of nicotine. The tobacco product is selected from cigarettes,
CC cigars, pipe tobacco, snuff, chewing tobacco, products blended with the
CC tobacco product and their mixtures. The nucleic acid molecule is useful
CC for altering plant phenotypes, thus producing a transgenic plant having
CC reduced levels of nicotine. This sequence represents cdNA encoding a
CC tobacco cytochrome P450 enzyme of the invention.
XX
SQ Sequence 1673 BP; 494 A; 309 C; 356 G; 514 T; 0 U; 0 Other;

Query Match 23.1%; Score 400.4; DB 14; Length 1673;
Best Local Similarity 57.7%; Pred. No. 4.4e-91;
Matches 810; Conservative 0; Mismatches 561; Indels 33; Gaps 4;

QY 113 CCTCAGGGCCATGAAGTTTCTATCATAGTAATCTTCTCATTTTATTACTCACTTCT 172
Db 135 CCACCAGGTCCATGGAACTACCAATACTAGGAAGTATGCTTCATATGCT- ----TGCT 188
QY 173 GATCTAGGCCATGAACGTTTATAGAGCCTTGGCTCAAATTTATGGACCTGTTATGCTCT 232
Db 189 GGACTACACACCATTGCTCTTAGAGATTAGCCAAAANAATATGGACCACTTATGACCTT 248
QY 233 CAAATTGGCCAGTTTTCAGCTGTTGCTCATTTCTTCAGCTGAAGCAGCCAAAGAGGTTATG 292
Db 249 CAAATTAGGTGAAGTTTCTGCGTTCGTGTTACTTCTCTGATACGCAAAAGAGTATTA 308
QY 293 AAAACTCAGGCTGATGCTTCCCAACGCCCTATCTGCTTCGAGCGCACAGATTGTTGTTT 352
Db 309 AAAACTCATGACATCGCTTTTTCGCTTAGGCTTAGCCTTTTGGCCCCCGAGATTGCTCTG 368
QY 353 TATAATCGGAAGATGCTTGTGTTCTCATATGGAGATCACTGGAGGCGAGATGAAGAA 412
Db 369 TACAATAGTCTGATCTAGCCTTTTGGCCCTATGGGACTATTGGAGACAAATGCGTAAA 428
QY 413 ATTTGGATACITTGAAATTTCTGAGTGCACAAAAGTTCAATCTCCAGGTTAATCCGAGAG 472
Db 429 ATATGCTCTTGGAAAGTCTCAGTGCAGAAATGTTGGACATTTTAGCTCTATTAGGCGG 488
QY 473 GAAGAAATGGAGGATGCCATCACATTTCTCGTTGAAAGCCGATCTCCCGTCAATATT 532
Db 489 AATGAAGTTCTTCTCGTCTCATTAATTTTATCCGGTCTATCTTCTGGTGAACCTATTATGTT 548

QY 533 ACAAGATCATTTATGGCATTATAATTTTCGATCATGATAAGAAACATCCGTTGGTAA---T 589
|||
Db 549 ACGGAAAGGATCTTTTGTTCACAGCTCATGACATGTAGATCAGCGTTTGGCGAAGTG 608
|||
QY 590 TGTAGCAAAAGAAAGATTGCTGAGTGTGCCGATGCGATGCAATGAGGCGGACGAGT 649
|||
Db 609 TTCAAAGAGCAAGACAAATTTATACAACTAATTAAGAAGTGATACTCTTAGCAGGAGGG 668
|||
QY 650 TTTGGCACCGCAGACGCTTTCCGACGTGGAAATTACTTCACTATATCATTTGAGCTGAG 709
|||
Db 669 TTTGATGGCTGACATATTCCTTCACTGAAGTTTCTTCATGTGCTCAGTGAATGAAG 728
|||
QY 710 TCAAAACCCGAGCGTTTGCATCAGGAGATTGACGATATACCTTTGAAGAGATTCTTAATGAA 769
|||
Db 729 GGTAGATTATGATGACACCATTAAGGTAGATGCCATTGTTGAGAATGTCATCAATGAG 788
|||
QY 770 CAC-----AAAGCCAATAGCCCTTTTGAAGCGGATCACTTAATG 808
|||
Db 789 CACAAGAAAATCTTGGCAATTGGGAAACTAATGGAGCGTTAGGAGGTGAAGATTTAAT 848
|||
QY 809 GATGTTCTATTGAATCTTCAAAAAAATGGAACGTTTCCAGTGCCAGTGACACAAACGAAAGC 868
|||
Db 849 GATGTTCTTAAGACTTATGAATGATGGAGGCTTCAATTTCTTATCACCACGACCAAC 908
|||
QY 869 ATCAAAGCATCCGTTTGTGCAAAATGTTTACTGCGGGAGCGGAAACAACTTCGAAAGCTACA 928
|||
Db 909 ATCAAAGCCATAATTTTGACATGTTTGTGCGGGACAGAGACTTCATCGTCAACAATT 968
|||
QY 929 GAATGGGTAATGCGAGAGCTGATGAAAATCCAACTGAACTAAGAAAGCACAAAGAA 988
|||
Db 969 GTGTGGGCTATGTTAGAAATGGTGAATAATCCAGCCGTATTCGCGAAAGCTCAAGCAGAA 1028
|||
QY 989 GTTAGACAAGTATTTGGTGAAATGGGAAAGTTGATGAATCAAGATTTTCATGATTTGAAA 1048
|||
Db 1029 GTTAGAGAGCATTTAGAGGAAAGAAACTTTTCGATGAAATGATGTGGAGGAGCTAAAC 1088
|||
QY 1049 TTTCTCAAGTTAGTGGTTAAAGAACTCTAAGATTACATCTCCGGTT---GTCTTGATT 1105
|||
Db 1089 TACCTAAAGTTAGTAATAAAGAAACTCTAAGACTTCATCCACCGGTTCCACTTTTGCTC 1148
|||
QY 1106 CCGAGGGAGTGTAGAGAAACACAGATTGATGGATGNAATTCATCCGAAACACTCGA 1165
|||
Db 1149 CCAAGAGAAATGTAGGAAGAGACAAATATAAACGGCTACACTATTCCTGTAAGACCANA 1208
|||
QY 1166 ATTGTTGTGAATGCTTTGGCGATAGGAGAGATCCTTAATACCTTGGTCGGAACTCGAAAG 1225
|||
Db 1209 GTCATGGTTAATGTTTGGGCTTTGGAGAGATCCAAAATA TTGGAATGACGCAGAACT 1268
|||
QY 1226 TTTAACCAGAAAGGTTTAAAGATTGTGCAATTGATTATTAAGGAGACGACATTTGAACTG 1285
|||
Db 1269 TTTATGCCAGAGAGATTGAGCAGTGCTCTAAGGATTTTGTGGTAATAATTTTGAATAT 1328
|||
QY 1286 GTACCATTTGGTGACGAAAAAGAAATATGCTCTGCGATTACTTCAGCTATTACCAATTG 1345
|||
Db 1329 CTTCCATTTGGTGGCGAAGGAGGATTGTCTCGGATTTCTGTTGGCTTAGCTAATGCT 1388
|||
QY 1346 GAGTATGTCATTATAAATCTATTATATCAATTTTAATTGGGAACTGGCCGATCGAATTACA 1405
|||
Db 1389 TATTTGCCATTGGCTCAATTACTATATCACTTCGATTGGAAACTCCCTGCTGGAATCGAA 1448
|||
QY 1406 CCTCAACACTTGTATATGATGAGCTATTGGGGTGCTCTCAGGAAAAAATAGATCTT 1465
|||
Db 1449 CCAAGCGACTTGGACTTGACTGAGTTGGTTGGATTAAC TGCCGCTAGAAAAAAGTGACCTT 1508
|||
QY 1466 AAGTTGATTCCTATTCCATATCAA 1489
|||
Db 1509 TACTTGGTTCGCACTCCTTATCAA 1532
|||

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 31, 2006, 09:33:07 ; Search time 2310 Seconds
(without alignments)
9218.386 Million cell updates/sec

Title: US-10-759-813-1
Perfect score: 1733
Sequence: 1 gataaaaggaaatggagc.....tcttttcaaatccgaaaaa 1733

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 18992170 seqs, 6143817638 residues
Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA Main:*

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
- 10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
- 11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
- 12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
- 13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
- 14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq:*
- 15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq:*
- 16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1733	100.0	1733	3	US-09-909-566C-1
2	1733	100.0	1733	8	US-10-759-813-1
3	410	23.7	1538	8	US-10-686-947-153
4	410	23.7	1538	10	US-10-943-507-153
5	410	23.7	1538	11	US-10-934-944-153
6	410	23.7	1538	16	US-11-116-881A-162
7	400.4	23.1	1673	8	US-10-686-947-209
8	400.4	23.1	1673	8	US-10-686-947-211
9	400.4	23.1	1673	8	US-10-686-947-231
10	400.4	23.1	1673	10	US-10-943-507-206
11	400.4	23.1	1673	10	US-10-943-507-208
12	400.4	23.1	1673	11	US-10-934-944-209
13	400.4	23.1	1673	11	US-10-934-944-211
14	400.4	23.1	1673	11	US-10-934-944-231
15	400.4	23.1	1673	16	US-11-116-881A-218
16	400.4	23.1	1673	16	US-11-116-881A-220
17	400.4	23.1	1673	16	US-11-116-881A-240

18	398.8	23.0	1673	8	US-10-686-947-213	Sequence 213, App
19	398.8	23.0	1673	10	US-10-943-507-210	Sequence 210, App
20	398.8	23.0	1673	11	US-10-934-944-213	Sequence 213, App
21	398.8	23.0	1673	16	US-11-116-881A-222	Sequence 222, App
22	393.2	22.7	1610	8	US-10-686-947-197	Sequence 197, App
23	393.2	22.7	1610	8	US-10-686-947-261	Sequence 261, App
24	393.2	22.7	1610	10	US-10-943-507-194	Sequence 194, App
25	393.2	22.7	1610	10	US-10-943-507-256	Sequence 256, App
26	393.2	22.7	1610	11	US-10-934-944-197	Sequence 197, App
27	393.2	22.7	1610	11	US-10-934-944-261	Sequence 261, App
28	393.2	22.7	1610	16	US-11-116-881A-206	Sequence 206, App
29	393.2	22.7	1610	16	US-11-116-881A-270	Sequence 270, App
30	388.8	22.4	1566	8	US-10-686-947-191	Sequence 191, App
31	388.8	22.4	1566	10	US-10-943-507-188	Sequence 188, App
32	388.8	22.4	1566	11	US-10-934-944-191	Sequence 191, App
33	388.8	22.4	1566	16	US-11-116-881A-200	Sequence 200, App
34	388.8	22.4	1576	8	US-10-686-947-149	Sequence 149, App
35	388.8	22.4	1576	10	US-10-943-507-149	Sequence 149, App
36	388.8	22.4	1576	11	US-10-934-944-149	Sequence 149, App
37	388.8	22.4	1576	16	US-11-116-881A-158	Sequence 158, App
38	388.8	22.4	1581	8	US-10-686-947-151	Sequence 151, App
39	388.8	22.4	1581	10	US-10-943-507-151	Sequence 151, App
40	388.8	22.4	1581	11	US-10-934-944-151	Sequence 151, App
41	388.8	22.4	1581	16	US-11-116-881A-160	Sequence 160, App
42	388.8	22.4	1664	8	US-10-686-947-199	Sequence 199, App
43	388.8	22.4	1664	10	US-10-943-507-196	Sequence 196, App
44	388.8	22.4	1664	11	US-10-934-944-199	Sequence 199, App
45	388.8	22.4	1664	16	US-11-116-881A-208	Sequence 208, App

ALIGNMENTS

RESULT 1

US-09-909-566C-1
; Sequence 1, Application US/09909566C
; Publication No. US20030066103A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Edgar B
; TITLE OF INVENTION: A Cytochrome P450 enzyme associated with the synthesis of delta-
; FILE REFERENCE: BBI465 US NA
; CURRENT APPLICATION NUMBER: US/09/909,566C
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 60/219833
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1733
; TYPE: DNA
; ORGANISM: Euphorbia lagascae
US-09-909-566C-1

Query Match	100.0%	Score 1733;	DB 3;	Length 1733;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1733;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GCATAAAGGAAATGGAGCAGAAAAATCTCTTTTCGAGCATTTTAAATAGTTTCT	60	
Db	1	GCATAAAGGAAATGGAGCAGAAAAATCTCTTTTCGAGCATTTTAAATAGTTTCT	60	
QY	61	GCTTCTTTTATCTTAGTAGTATCATGAGGTTGGAGAGAACAGAAATCCACCTCCAG	120	
Db	61	GCTTCTTTTATCTTAGTAGTATCATGAGGTTGGAGAGAACAGAAATCCACCTCCAG	120	
QY	121	GCATGGAAGTTTCTTATCATAGGTAATCTTCTCATTTTATTTACTTCATCTGATCTAGG	180	
Db	121	GCATGGAAGTTTCTTATCATAGGTAATCTTCTCATTTTATTTACTTCATCTGATCTAGG	180	
QY	181	CCATGAACGTTTTAGAGCCTTGGCTCAAAATTTATGGACCTGTATGAGTCTTCAAAATGG	240	
Db	181	CCATGAACGTTTTAGAGCCTTGGCTCAAAATTTATGGACCTGTATGAGTCTTCAAAATGG	240	

```
QY 241 CCAAGTTTCAGCTGTGTGTCATTTCTTCAGCTGAAGCAGCAAGAGGTTATGAAAACTCA 300
DB |||||||
DB 241 CCAAGTTTCAGCTGTGTGTCATTTCTTCAGCTGAAGCAGCAAGAGGTTATGAAAACTCA 300
QY 301 GGCTGATGCTTCGGCCCAACGCCCTATCGTCTTGGACGCAACAGATGTGTTTTATATCG 360
DB |||||||
DB 301 GGCTGATGCTTCGGCCCAACGCCCTATCGTCTTGGACGCAACAGATGTGTTTTATATCG 360
QY 361 GAAAGATGCTGTGTTGCTTCATATGAGATCAGCTGGAGGCGAGATGAAGAAAAATTTGGAT 420
DB |||||||
DB 361 GAAAGATGCTGTGTTGCTTCATATGAGATCAGCTGGAGGCGAGATGAAGAAAAATTTGGAT 420
QY 421 ACTTGAATTTCTGAGTGCCAAAAAGTTCAATCTCCAGGTTAAATCCGAGAGGAGAAAT 480
DB |||||||
DB 421 ACTTGAATTTCTGAGTGCCAAAAAGTTCAATCTCCAGGTTAAATCCGAGAGGAGAAAT 480
QY 481 GGAGATGCGCATCAATTCCTCGTTGAAAGCCGGATCTCCGGTCAATATTAACAAGAT 540
DB |||||||
DB 481 GGAGATGCGCATCAATTCCTCGTTGAAAGCCGGATCTCCGGTCAATATTAACAAGAT 540
QY 541 CATTATATGCGATTATAATTTGCGATCATATGAAGACATCCGTTGGTAAATTTGAAGCAAAA 600
DB |||||||
DB 541 CATTATATGCGATTATAATTTGCGATCATATGAAGACATCCGTTGGTAAATTTGAAGCAAAA 600
QY 601 AGAAAGATTGCTGAGTGTGCGGATGCGAGTCAATGAGGCGAGCGAGTTTGGCACCGC 660
DB |||||||
DB 601 AGAAAGATTGCTGAGTGTGCGGATGCGAGTCAATGAGGCGAGCGAGTTTGGCACCGC 660
QY 661 AGACGCTTTTCGAGCGTGAATTAATCTTCACTATATCATTTGGAGCTGAGTCMAAACCCAG 720
DB |||||||
DB 661 AGACGCTTTTCGAGCGTGAATTAATCTTCACTATATCATTTGGAGCTGAGTCMAAACCCAG 720
QY 721 GCGTTTGATCAGGAGATTGAGATATCTTGAAGAGATCTTGAATGAACACAAAGCCAA 780
DB |||||||
DB 721 GCGTTTGATCAGGAGATTGAGATATCTTGAAGAGATCTTGAATGAACACAAAGCCAA 780
QY 781 TAAGCCTTTTGAAGCGGATACTTAATGATGTTCTATTTGAATCTTCAAAAAAATGGAAA 840
DB |||||||
DB 781 TAAGCCTTTTGAAGCGGATACTTAATGATGTTCTATTTGAATCTTCAAAAAAATGGAAA 840
QY 841 CGTTCAGTCCAGTGACAAAGCAAGCATCAAGCATCCGTTTTCGCAATGTTTACTGC 900
DB |||||||
DB 841 CGTTCAGTCCAGTGACAAAGCAAGCATCAAGCATCCGTTTTCGCAATGTTTACTGC 900
QY 901 CGGGAGCGAAACAACTTCGAAAGCTACAGAAATGGGTAAATGGCAGAGCTGATGAAAAATCC 960
DB |||||||
DB 901 CGGGAGCGAAACAACTTCGAAAGCTACAGAAATGGGTAAATGGCAGAGCTGATGAAAAATCC 960
QY 961 AACTGAACCTAAGAAAAGCAACAAGAAAGTTAGACAAGTATTTGGTGAATGGGAAAAGT 1020
DB |||||||
DB 961 AACTGAACCTAAGAAAAGCAACAAGAAAGTTAGACAAGTATTTGGTGAATGGGAAAAGT 1020
QY 1021 TGATGAATCAAGATTTCAATTTGAAATTTTCAAGTTAGTGGTTAAGAACTCTAAG 1080
DB |||||||
DB 1021 TGATGAATCAAGATTTCAATTTGAAATTTTCAAGTTAGTGGTTAAGAACTCTAAG 1080
QY 1081 ATTACATCTCCGGTTGCTTCAATTCGAGGAGGTGAGAGAAACAAACAGAAATTTGATGG 1140
DB |||||||
DB 1081 ATTACATCTCCGGTTGCTTCAATTCGAGGAGGTGAGAGAAACAAACAGAAATTTGATGG 1140
QY 1141 ATATGAAATTTATCCGAACACTCGAATTTGTTGTAATGCTTCGGCGATAGGAAGATCC 1200
DB |||||||
DB 1141 ATATGAAATTTATCCGAACACTCGAATTTGTTGTAATGCTTCGGCGATAGGAAGATCC 1200
QY 1201 TAATACTTGGTCGGAACCTGGAAGTTTAACCCAGAAAGTTTAAGATTGCGAATGA 1260
DB |||||||
DB 1201 TAATACTTGGTCGGAACCTGGAAGTTTAACCCAGAAAGTTTAAGATTGCGAATGA 1260
QY 1261 TTATAAAGGAGACGATTTGAATCTGATACATTTTGGTCAGGAAAAAGAAATATGCTCTGG 1320
DB |||||||
DB 1261 TTATAAAGGAGACGATTTGAATCTGATACATTTTGGTCAGGAAAAAGAAATATGCTCTGG 1320
```

```
QY 1321 CATTACTTCAGCTATTACCAATTTTGGAGTATGTCAATTTATAAATCTATATATCATTTTAA 1380
DB |||||||
DB 1321 CATTACTTCAGCTATTACCAATTTTGGAGTATGTCAATTTATAAATCTATATATCATTTTAA 1380
QY 1381 TTGGGAACCTGCCCGATGGAATTTACACCTCAAACTCTGATATGACTGAAGCTATTGGCGG 1440
DB |||||||
DB 1381 TTGGGAACCTGCCCGATGGAATTTACACCTCAAACTCTGATATGACTGAAGCTATTGGCGG 1440
QY 1441 TGCTCTCAGGAAAAAATAGATCTTTAAGTTGATTTCTTATTCATATCAAGTTAGCTTAGG 1500
DB |||||||
DB 1441 TGCTCTCAGGAAAAAATAGATCTTTAAGTTGATTTCTTATTCATATCAAGTTAGCTTAGG 1500
QY 1501 CTCAAAATATTCTTGAATTAATAGAGGTTGAAATATATATATAATAAATCTTTAATTAAAG 1560
DB |||||||
DB 1501 CTCAAAATATTCTTGAATTAATAGAGGTTGAAATATATATAATAAATCTTTAATTAAAG 1560
QY 1561 ATGTTCTTAATATGTTTGGGTGAGTTAATAGGTTTTCCACCGCATATATAAGTAGCCT 1620
DB |||||||
DB 1561 ATGTTCTTAATATGTTTGGGTGAGTTAATAGGTTTTCCACCGCATATATAAGTAGCCT 1620
QY 1621 TCCTTGATGATGGGTTAGATTATAATAGTTGTGGGTTGGATTTTATAGATGGGTTAAAT 1680
DB |||||||
DB 1621 TCCTTGATGATGGGTTAGATTATAATAGTTGTGGGTTGGATTTTATAGATGGGTTAAAT 1680
QY 1681 GATTTGGATGATTAATAATAAATTTGAAATGTTTTCTTTTCCAAATCCGAAAAA 1733
DB |||||||
DB 1681 GATTTGGATGATTAATAATAAATTTGAAATGTTTTCTTTTCCAAATCCGAAAAA 1733

RESULT 2
US-10-759-813-1
; Sequence 1, Application US/10759813
; Publication No. US20040139499A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Edgar B
; TITLE OF INVENTION: A Cytochrome P450 enzyme associated with the synthesis of delta-
; FILE REFERENCE: BBI465 US NA
; CURRENT APPLICATION NUMBER: US/10/759,813
; CURRENT FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: US/09/909,566C
; PRIOR FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 60/219833
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1733
; TYPE: DNA
; ORGANISM: Euphorbia lagascae
US-10-759-813-1

Query Match 100.0%; Score 1733; DB 8; Length 1733;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1733; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCATAAAAGGAAAATGGAGCAGAAAATCTCTCTTTCCGAGCATTTTAATAGTTTCT 60
DB 1 GCATAAAAGGAAAATGGAGCAGAAAATCTCTCTTTCCGAGCATTTTAATAGTTTCT 60
QY 61 GCTTGTGTTTATCTTAGTAGTAGTCATGAGGTGTGGAGAAACAGAAATCCACCTCCAGG 120
DB 61 GCTTGTGTTTATCTTAGTAGTAGTCATGAGGTGTGGAGAAACAGAAATCCACCTCCAGG 120
QY 121 GCCATGGAAGTTTCCATCATAGGTAATCTTCTCATTTATTACTCAGTTCTGTAGCTAGG 180
DB 121 GCCATGGAAGTTTCCATCATAGGTAATCTTCTCATTTATTACTCAGTTCTGTAGCTAGG 180
QY 181 CCATGAACGTTTTAGAGCCTTGGCTCAAAATTTATGGACCTGTGTATGAGTCTTCAAAATGG 240
DB 181 CCATGAACGTTTTAGAGCCTTGGCTCAAAATTTATGGACCTGTGTATGAGTCTTCAAAATGG 240
QY 241 CCAAGTTTCAGCTGTGTGTCATTTCTTCAGCTGAAGCAGCAAGAGGTTATGAAAACTCA 300
```

Db 241 CCAAGTTTCAGCTGTGTGTCATTTCTTCAGCTGAAGCAGCAAGAGGTTATGAAACTCA 300
Qy |||||
Db 301 GCGTATGCTTCGCGCAAGCCCTATCGTCTTGGACGACACAGATTGTGTTTATATCG 360
Qy |||||
Db 301 GCGTATGCTTCGCGCAAGCCCTATCGTCTTGGACGACACAGATTGTGTTTATATCG 360
Qy |||||
Db 361 GAAAGATGTCCTGTTGCTTCATATGAGATCACTGGAGCGAGATGAAGAAAATTTGGAT 420
Qy |||||
Db 361 GAAAGATGTCCTGTTGCTTCATATGAGATCACTGGAGCGAGATGAAGAAAATTTGGAT 420
Qy |||||
Db 421 ACTTGAATTTCTGAGTGCACAAAAGTTCAATCTCCAGGTTAATCCGAGAGAAAT 480
Qy |||||
Db 421 ACTTGAATTTCTGAGTGCACAAAAGTTCAATCTCCAGGTTAATCCGAGAGAAAT 480
Qy |||||
Db 481 GGAGGATGTCATCACATTCCTCGTTTCAAAGCCGATCTCCGGTCAATATTACAAGAT 540
Qy |||||
Db 481 GGAGGATGTCATCACATTCCTCGTTTCAAAGCCGATCTCCGGTCAATATTACAAGAT 540
Qy |||||
Db 541 CATTTATGCGATTTAATTTTCGATCATGATAAGAACATCCGTTGTTAAATTTGAAGCAAAA 600
Qy |||||
Db 541 CATTTATGCGATTTAATTTTCGATCATGATAAGAACATCCGTTGTTAAATTTGAAGCAAAA 600
Qy |||||
Db 601 AGAAGATTTCTGAGTGTTCGCGATGCAATGAGGCGAGCGAATTTGGCACCCG 660
Qy |||||
Db 601 AGAAGATTTCTGAGTGTTCGCGATGCAATGAGGCGAGCGAATTTGGCACCCG 660
Qy |||||
Db 661 AGACGCTTTTCCGAGCGTGAATTTACTTCACTATATCAATGGAGCTGATCAAAACCCAG 720
Qy |||||
Db 661 AGACGCTTTTCCGAGCGTGAATTTACTTCACTATATCAATGGAGCTGATCAAAACCCAG 720
Qy |||||
Db 721 GCGTTTGCATCAGGAGATTGACGATATACCTTGAAGAGATTTCTTAATGAACACAAAGCCAA 780
Qy |||||
Db 721 GCGTTTGCATCAGGAGATTGACGATATACCTTGAAGAGATTTCTTAATGAACACAAAGCCAA 780
Qy |||||
Db 781 TAAGCCTTTTGAAGCGGATACTTAATGATGTTCTTAATGAATCTTCAAAAAAATGGA 840
Qy |||||
Db 781 TAAGCCTTTTGAAGCGGATACTTAATGATGTTCTTAATGAATCTTCAAAAAAATGGA 840
Qy |||||
Db 841 CGTTCAGTGCAGTGACAAAAGCAAGAGCATCAAGCATCCGTTTTCGCAATGTTTACTGC 900
Qy |||||
Db 841 CGTTCAGTGCAGTGACAAAAGCAAGAGCATCAAGCATCCGTTTTCGCAATGTTTACTGC 900
Qy |||||
Db 901 CGGAGCGCAAAACAACTTTCGAAAGCTACAGAATGGGTAAATGGCAGAGCTGATGAAAATCC 960
Qy |||||
Db 901 CGGAGCGCAAAACAACTTTCGAAAGCTACAGAATGGGTAAATGGCAGAGCTGATGAAAATCC 960
Qy |||||
Db 961 AACTGAATTAAGAAAAGCAAGAAAGATTAGACAAAGTATTTGGTGAATCGGAAAAGT 1020
Qy |||||
Db 961 AACTGAATTAAGAAAAGCAAGAAAGATTAGACAAAGTATTTGGTGAATCGGAAAAGT 1020
Qy |||||
Db 1021 TGATGAATCAAGATTTCAATGTTTGAATTTGAAAATTTCTCAAGTTAGTGGTTAAAGAACTTAAG 1080
Qy |||||
Db 1021 TGATGAATCAAGATTTCAATGTTTGAATTTGAAAATTTCTCAAGTTAGTGGTTAAAGAACTTAAG 1080
Qy |||||
Db 1081 ATTACATCTCCGGTGTCTTCATTTCCGAGGAGTGTAGAGAAAACAACGAATTTGATGG 1140
Qy |||||
Db 1081 ATTACATCTCCGGTGTCTTCATTTCCGAGGAGTGTAGAGAAAACAACGAATTTGATGG 1140
Qy |||||
Db 1141 ATATGAAATTTCAATCCGAAACACTTCGAATTTGTTGTAATGTTGGGCGATAGGAAGATCC 1200
Qy |||||
Db 1141 ATATGAAATTTCAATCCGAAACACTTCGAATTTGTTGTAATGTTGGGCGATAGGAAGATCC 1200
Qy |||||
Db 1201 TAATACTTGGTCCGAACTTGAAGTTTAAACCCAGAAAGTTTAAAGATTGTCGAATTTGA 1260
Qy |||||
Db 1201 TAATACTTGGTCCGAACTTGAAGTTTAAACCCAGAAAGTTTAAAGATTGTCGAATTTGA 1260
Qy |||||
Db 1261 TTATAAGGAGACGATTTGAATTCGGTACCATTTGGTGCAGGAAAAGAAATATGTCCTGG 1320
Qy |||||
Db 1261 TTATAAGGAGACGATTTGAATTCGGTACCATTTGGTGCAGGAAAAGAAATATGTCCTGG 1320
Qy |||||
Db 1321 CATTAATTCAGCTATTACAAATTTGGAGTATGTCATTAATAATCTTATATATCATTTTAA 1380
Qy |||||

Db 1321 CATTAATTCAGCTATTACCAATTTGGAGTATGTCAATTAATAAATCTATTATATCATTTTAA 1380
Qy |||||
Db 1381 TTGGGAACCTGGCGATGGAAATTACACCTCAACACATTTGATATGACATGACATTTGCGCG 1440
Qy |||||
Db 1381 TTGGGAACCTGGCGATGGAAATTACACCTCAACACATTTGATATGACATGACATTTGCGCG 1440
Qy |||||
Db 1441 TGCTCTCAGGAAAAAATAGATCTTTAAAGTTGATTCCTATTTCATATCAAGTTAGCTTAGG 1500
Qy |||||
Db 1441 TGCTCTCAGGAAAAAATAGATCTTTAAAGTTGATTCCTATTTCATATCAAGTTAGCTTAGG 1500
Qy |||||
Db 1501 CTCAAATATTTCTTGATACATAGAGGGTTGAAATATATATATATAAATAACTTTTAAACG 1560
Qy |||||
Db 1501 CTCAAATATTTCTTGATACATAGAGGGTTGAAATATATATATAAATAACTTTTAAACG 1560
Qy |||||
Db 1561 ATGTTCTATATATGTTGGTGGATTAATAGTTTTCACCGATCATATAAGTAGCCT 1620
Qy |||||
Db 1561 ATGTTCTATATATGTTGGTGGATTAATAGTTTTCACCGATCATATAAGTAGCCT 1620
Qy |||||
Db 1621 TCTTTGATGATGGTTAGATTATATAGATTGTGGTTGGATTTTAGATGGTTAAAT 1680
Qy |||||
Db 1621 TCTTTGATGATGGTTAGATTATATAGATTGTGGTTGGATTTTAGATGGTTAAAT 1680
Qy |||||
Db 1681 GATTTGGATGATATAATAAATGAAATGTTTCTTTTCAAAATCCGAAAAA 1733
Qy |||||
Db 1681 GATTTGGATGATATAATAAATGAAATGTTTCTTTTCAAAATCCGAAAAA 1733

RESULT 3

US-10-686-947-153
; Sequence 153, Application US/10686947
; Publication No. US20040162420A1
; GENERAL INFORMATION:
; APPLICANT: Profigen Inc.
; TITLE OF INVENTION: Cloning of Cytochrome P450 Genes from Tobacco
; FILE REFERENCE: 79601
; CURRENT APPLICATION NUMBER: US/10/686,947
; CURRENT FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: US 10/387346
; PRIOR FILING DATE: 2003-03-12
; NUMBER OF SEQ ID NOS: 298
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 153
; LENGTH: 1538
; TYPE: DNA
; ORGANISM: NICOTIANATABACUM
US-10-686-947-153

Query Match 23.7%; Score 410; DB 8; Length 1538;
Best Local Similarity 58.1%; Pred. No. 4.4e-91;
Matches 816; Conservative 0; Mismatches 555; Indels 33; Gaps 4;

Qy 113 CCTCAGGCGCCATGGAAAGTTTCCCTATCATAGTATCTTCCTCATTTTACTCATTCT 172
Db |||||
129 CCACAGGTCCATGGAAACTACCAATACCTAGGAAGTATGCTTCATATGTT-----TGGT 182
Qy 173 GATCTAGGCCATGAACGTTTGTAGAGCTTTGAGCGCTTGGCTCAAAATTTATGAGCTCTT 232
Db |||||
183 GGACTACACACCATGTCCTTAGAGATTAGCCAAAATAATGAGCCGCTTAGCACCTT 242
Qy 233 CAAATGGCCCAAGTTTTCAGCTGTTGTCATTTCTTCAGCTGAAAGAGCCAGCAAGAGTTATG 292
Db |||||
243 CAAATAGGTGAAGTTTCTGCAAGTTGTGTTTACTTCTCTGATATATGGCAAAAGAGTACTA 302
Qy 293 AAAACTCAGGCTGATGCTTCCGCCAAACGCCCTATCGTCTTGAGCGCACAGATTGTGTT 352
Db |||||
303 AAAACTCATGATCATCGCTTTCGCGTCTAGGCCCTAGCCCTTTTGGCCCCGGAGATTGCTGT 362
Qy 353 TATAATCGGAAAGATGCTGTTGTTTCTCATATGAGATCACTGGAGGAGAGATGAAGAAA 412
Db |||||
363 TACAATAGGTCTGATCTTGGCGTTTGGCCCTTATGCGGATTTATGGAGACAAATGCGTAAA 422
Qy 413 ATTTGGATATCTGAATTTCTGAGTGCCCAAAAAAGTTTCAATCTCAGGTTTAAATCCGAGAG 472
Db |||||

Db 423 ATATGTGCTTGGAAAGTCTCAGTGCCAAAGAAATGTTCCGACATATAGTCTATTAGGCGC 482
Qy 473 GAAGAAATGGAGGATGCCATCACATTTCTCGGTTCGAAAGCGGATCTCCGGTCAATATT 532
Db 483 GATGAAGTCTTCGTCTCTTAATTTTATCCGGTCATCTTCTGGTGAGCCTGTTAATATT 542
Qy 533 ACAAGATCATTTATGGCATTATAATTTCCGATCATGATAAGAACATCCGTTGGTAA---T 589
Db 543 ACGGAAAGGATCTTTTGTTCACAAAGTCCATGACATGTAGATCAGCTTTGGGCAAGTA 602
Qy 590 TGTAAACAAAGAAAGATGCTGAGTGTGCCGATGCAGTCAATGAGCGGACGACGAGT 649
Db 603 TTCAAGGAGCAAGACAAATTTATACAACTAATTAAGAAAGTTATACTCTTAGCAGGAGG 662
Qy 650 TTTGCGACCGCAGACGCTTTTCCGACGTGGAAATTTACTTCACTATATCATTTGGAGCTGAG 709
Db 663 TTTGATGTGGCTGACATATTCCTTCATCAAGTCTCTTCATGTCTCAGTGGATGAAG 722
Qy 710 TCAAAACCCAGCGGTTTGCATCAGAGATGTAGCATATACTTGAAGAGATTTCTTAATGAA 769
Db 723 GGTAAATATGAATGCACACCATAAAGGTAGATGCTATTGTTGAGAATGTCAACAACGAG 782
Qy 770 CAC-----NAAGCCATAGCCTTTTGAAGCGGATAACTTTAATG 808
Db 783 CACAAGAAAAATCTTGCAATTTGGGAAACTAATGAGCGTTAGGAGGTGAAGATTTAATT 842
Qy 809 GATGTTCTATTGAATCTTCAAAAAATGGAACGTTCCAGTGCCAGTGCACAAACGAAAGC 868
Db 843 GATGTTCTTAAACTTATGATGATGAGGCGCTTCAATTTCTCTTACCAACGACAAAC 902
Qy 869 ATCAAGCATCGTTTGGCAATGTTTACTGCGGAGCGAAACAACTTCGAAAGCTACA 928
Db 903 ATCAAGCTATAATCTTTGACATGTTTGTCTGGAACAGAGACTTCATCGTCAACAATT 962
Qy 929 GAATGGGTAAAGCGAGCTGATGAATAATCAACTGAACCTAAGAAAGCAACAAGAA 988
Db 963 GTGTGGGCTATGGTGAATGTTGAAATAATCCAATGTTATTTGCGAAAGCTCAACGAA 1022
Qy 989 GTTAGCAAGTATTTGGTGAATGGGAAAGTTGATGAATCAAGATTTTCATGATTGAAA 1048
Db 1023 GTAAGAGATGATTTAGAGAAAGAAAGAACTTTTGATGAATAATGATGGAGGAGCTAAAC 1082
Qy 1049 TTCCTCAAGTTAGTGGTTAAAGAACTCTAAGATTTACATCTCCGGTT---GTCCTGATT 1105
Db 1083 TATCTAAAGTTAGTCAATTAAGAAACTCTAAGACTTTCATCCACCGGTTCCACTTTTGCTC 1142
Qy 1106 CCGAGGAGGTAGAGAAACACGAAATTTGATGATGAAATTCATCCGAACTCGA 1165
Db 1143 CCNAGAAATGTAGGAAAGAGACAAATATAACGGCTACACTATTCCTGTAAAGACCAA 1202
Qy 1166 ATTTGTTGAAATGCTTTGGCGGATAGGAAGAGATCCCTAATCTTTGTTGGAACTCGGAAAG 1225
Db 1203 GTCATGTTAATGTTTGGGCAATGGGAAGAGATCCAAATAATGGGATGATCGAGAACT 1262
Qy 1226 TTTAACCCGAAAGGTTTAAAGATTTGCAATTTGATTAATAAGGAGCAGCATTTGAACATG 1285
Db 1263 TTTAAGCCAGAGAGATTTGAGCAGTGTCTAAGGATTTTGTGTTGTAATAATTTGAATAT 1322
Qy 1286 GTACCAATTTGGTGCAGGAAAGAAATATGCTCTGGCATTACTTCAGCTATTACCAATTG 1345
Db 1323 CTTCAATTTGGTGGGAAAGGAGGATTTGTCCAGGAAATTCGTTGGTTTACCTAATGCT 1382
Qy 1346 GAGTATGTCATTATAAATCTATTATATCAATTTTAAATGGGAACCTGGCCGATGGAATTACA 1405
Db 1383 TATTTGCCATTTGCTCAATTAATCTTATCACTTTTATGTTGGAACTCCCACTGGAATCAA 1442
Qy 1406 CTTCAACAACCTGATATGACTGAAGCTATTTGGCGGTGCTCTCAGGAAAGAAATATGATCTT 1465
Db 1443 CCAAGCGACTTGGACTTGAATCTGAGTTGGTTGGAGTAACCTGCCGCTAGAAAAAAGTACCTT 1502
Qy 1466 AGTTGATTTCTTATTCATATCAA 1489
Db 1503 TACTTGGTGGCACTCTTATCAA 1526

RESULT 4

US-10-943-507-153
; Sequence 153, Application US/10943507
; Publication No. US2005013244A1
; GENERAL INFORMATION:
; APPLICANT: U.S. Smokeless Tobacco Company
; TITLE OF INVENTION: 83492-7270 CIP - Cloning of Cytochrome P450 Genes From Nicotiana
; FILE REFERENCE: 83492-7270
; CURRENT APPLICATION NUMBER: US/10/943,507
; CURRENT FILING DATE: 2004-09-17
; PRIOR APPLICATION NUMBER: 10/686,947
; PRIOR FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: 10/387,346
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: 10/340,861
; PRIOR FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 60/363,684
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/347,444
; PRIOR FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 60/337,684
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 153
; LENGTH: 1538
; TYPE: DNA
; ORGANISM: NICOTIANATABACUM
; US-10-943-507-153

Query Match 23.7%; Score 410; DB 10; Length 1538;
Best Local Similarity 58.1%; Pred. No. 4.4e-91;
Matches 816; Conservative 0; Mismatches 555; Indels 33; Gaps 4;

Qy 113 CCTCCAGGCGCCATGGAAAGTTTCCCTATCATAGTAACTCTTCCTCATTTCTCATTTACTCACTTCT 172
Db 129 CCACAGGTCCATGGAAACTACCAATACTGGAAGTATGCTTCATATGGT-----TGGT 182
Qy 173 GATCTAGGCCATGAACGTTTTTAGAGCCTTGGCTCAAAATTTATGGACCTGTTTATGAGTCTT 232
Db 183 GGACTACCAACCATGTCCTTAGAGATTTAGCCAAAAATATGGACCGCTTATGCACCTT 242
Qy 233 CAATTTGCCCAAGTTTCAGCTGTTGTCATTTCTTCAGCTGAAGCAGCACAAGAGGTTATG 292
Db 243 CAATTTAGTGAAAGTTTCTGCAGTTGTGGTTACTTCTCTGATATGCGCAAAAGAGTACTA 302
Qy 293 AAAAATCAGGCTGATGCTTCGCCCAACGCCCTATCGTCTTGGACGACAGATTTGTGTTT 352
Db 303 AAAATCTATGACATCGCTTTCGCTTAGCCCTAGCCCTTTGGCCCGCGAGATTTGTCTGT 362
Qy 353 TATAATCGGAAAGATGCTTGTGTTGCTTCATATGGAGATCACTGGAGGCGAGATGAAGAAA 412
Db 363 TACAATAGTCTGATCTTGGCGTTTGGCCCTATGGCGATTAATGGAGACAAATGCGTAAA 422
Qy 413 ATTTGGATCTTGAATTTCTGAGTGCCCAAAAAGTTCAATCTCCAGGTTAATCCGAGAG 472
Db 423 ATATGTGCTTGGAAAGTGTCTAGTGCCCAAGAAATGTTCCGACATATAGCTCTATTAGGCC 482
Qy 473 GAAGAAATGGAGGATGCCATCACATTTCTCGGTTCGAAAGCGGATCTCCGGTCAATATT 532
Db 483 GATGAAGTTCTTGTGCTCCTTAATTTTATCCGGTCACTTCTGTGGTGAGCTGTTAATATT 542
Qy 533 ACAAGATCATTTATGGCAATTATAATTTTCGATCATGATAAGAAACATCCGTTGGTAA---T 589
Db 543 ACGGAAAGGATCTTTTGTTCACAAGTCCATGACATGTAGATCAGCGTTTGGGCAAGTA 602
Qy 590 TGTAAAGCAAAAGAAAGATTTGCTGAGTGTTCGGATGCTCAATGAGGACGCGACGACT 649
Db 603 TTCAAGGAGCAGACAAATTTATACAACTAATTAAGAAAGTTTATACTCTTTAGCAGGAGG 662
Qy 650 TTTGCAACCGCAGACGCTTTTCCGACGTGGAATTTACTTCACTATATCATTTGAGGCTGAG 709


```
Db 663 TTTGATGTGGCTGACATATTTCCCTTCATACAAGTCTCTTCATGTGCTCAAGTGAATGAAG 722
Qy 710 TCAAAACCCAGGCGTTTGCATCAGGAGATTGACGATATACCTTGAAGAGATTCTTAATGAA 769
Db 723 GGTAAGATTATGAATGCACACCATAGAGGTAGATGCTATTGTTGAGAATGTCTATCAACGAG 782
Qy 770 CAC-----AAAGCCAAATAAGCCTTTTGAAGCGATAACTTAAATG 808
Db 783 CACAAGAAAATCTTCCAAATTCGGAAAATCTAATGGAGCGTTAGGAGGTGAAGTTTAAT 842
Qy 809 GATGTTCTATTGAATCTTCAAAAAATGGAAGCGTTCCAGTGCCAGTGACAAACGAAAGC 868
Db 843 GATGTTCTTCTAAACCTTGAATGATGAGGCGCTTCAATTTCTTATCACCACGACAAC 902
Qy 869 ATCAAGCATCCGTTTGTGCAATTTTACTGCGGAGGAGCAACAACTTCGAAGCTACA 928
Db 903 ATCAAGCATTAATCTTTGACATGTTTGTCTGGAACAGAGACTTCATCGTCAACAAT 962
Qy 929 GAATGGTAAATGGCAGAGCTGATGAAAAATCCAACTGAACTAAGAAAAACACAAGAA 988
Db 963 GTGTGGCTATGTTGGAATGTTGAAAATCCAACTGATTTTCCGNAAGCTCAAGCAGAA 1022
Qy 989 GTTAGCAAGTATTGGTGAAATGGGAAAAGTTGATGAATCAAGATTTTCATGATTTGAA 1048
Db 1023 GTAAGAGATGCAATTTAGAGAAAAAGAACTTTTGTATGAAAAATGATGTGGAGGAGCTAAAC 1082
Qy 1049 TTTCTCAAGTTAGTGTAAAGAACTCTAAGATTACATCTCCGGTT---GTCATTGATT 1105
Db 1083 TATCTAAAGTTAGTCTAATAAGAACTCTAAGACTTCATCCACCGGTTCCACTTTTGCTC 1142
Qy 1106 CCGAGGAGGTGATAGAAACAAACACGAAATGATGATATGAATTCATCCGAACACTCGA 1165
Db 1143 CCAAGAGATGTAGGAGAGACAAATATAACGGCTACACTATTCTGTAAAGACCAA 1202
Qy 1166 ATTGTTGTAATGCTTTGGCGATAGGAAGAGATCTTAATPACTTTGTCGGAACCTGGAAG 1225
Db 1203 GTCATGTTTAATGTTTGGGCAATTTGGAGAGATCCAAATATTTGGGATGATCAGAACT 1262
Qy 1226 TTTAACCCGAAGGTTTAAAGATGTGCAATTTGATTAATAAGGAGACATTTGNACTG 1285
Db 1263 TTTAAGCCAGAGAGATTTGAGCAGTCTCTAAGGATTTTGTGTTAAATATTTTGAATAT 1322
Qy 1286 GTACCAATTTGGTGAGGAAAAAGAAATATGTCCTGGCATTACTTCAGCTATTACCAATTTG 1345
Db 1323 CTTCACATTTGGTGGGAGGAGGATTTGTCCAGGATTTGTTGGTTTACCTAATGCT 1382
Qy 1346 GAGTATGCTATTATAAATCTATTATATCAATTTTAAATTTGGAACTGGCCGATGGAATTACA 1405
Db 1383 TATTTGCCATTGGCTCAATTAATCTTTTATCACTTTTGTATTTGGAACTCCCACTGGAATCAA 1442
Qy 1406 CCTCAACACTTGATATGACTGAAGCTATTGGCGGTGCTCTCAGGAAAAAATAGATCTT 1465
Db 1443 CCAAGCGACTTGGACTGACTGAGTTGGTGGAGTAACCTGCGCTAGAAAAAATGACCTT 1502
Qy 1466 AAGTTGATTCCTATTCCATATCAA 1489
Db 1503 TACTTGGTTGGACTTCCTTATCAA 1526
```

RESULT 5

```
US-10-934-944-153
; Sequence 153, Application US/10934944
; Publication No. US20060037096A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Dongmei
; TITLE OF INVENTION: Cloning of Cytochrome P450 Genes From Nicotiana
; FILE REFERENCE: 07678/141008
; CURRENT APPLICATION NUMBER: US/10/934,944
; CURRENT FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: 10/686,947
; PRIOR FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: 60/503,989
```

```
; PRIOR FILING DATE: 2003-09-18
; PRIOR APPLICATION NUMBER: 60/485,368
; PRIOR FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: 10/387,346
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: 10/340,861
; PRIOR FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 10/293,252
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 60/418,933
; PRIOR FILING DATE: 2002-10-16
; PRIOR APPLICATION NUMBER: 60/363,684
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/347,444
; PRIOR FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 60/337,684
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 387
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 153
; LENGTH: 1538
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
; US-10-934-944-153
```

```
Query Match 23.7%; Score 410; DB 11; Length 1538;
Best Local Similarity 58.1%; Pred. No. 4.4e-91;
Matches 816; Conservative 0; Mismatches 555; Indels 33; Gaps 4;

Qy 113 CTTCCAGGCGCCATGGAAGTTTCTCATCATAGTAGTAATCTTCTCTCATTTATTACTCACTTCT 172
Db 129 CCACCAGGTCCATGGAACACTACCAATACTAGGAAGTATGCTTCATATGCT-----TGGT 182
Qy 173 GATCTAGGCCATGAACGTTTTAGAGCCTTTGGCTCAAAATTTATGAGCCTGTTATGAGCTTT 232
Db 183 GGACTACCAACACCATGTCCTTAGAGATTTAGCCAAAAAATATGAGCCGCTTATGCACCTT 242
Qy 233 CAAATTTGCCAAGTTTTCAGCTGTTGCTCAATTTCTCAGCTGGAAGCAGCCAAAGGTTATG 292
Db 243 CAATTAGGTGAAGTTTTCGAGTGTGTTGCTTCTCTCTGATATGGCAAAAGTACTA 302
Qy 293 AAAAATCTAGGCTGATGCTCTTGGCCCAACGCTTATCGTCTTGGAGCGCAGAGTTGTGTTT 352
Db 303 AAAAATCTAGCATGCTCTTGGCGTCTAGGCTTAGCCCTTTGGCCCGGAGATTTGCTCT 362
Qy 353 TATAATCCGAAGAGTGTCTGTTTCTCATATGAGATCACTGGAGGAGATGAAGAAA 412
Db 363 TACAATAGGTCTGATCTTGGCGTTTGGCCCTATGGCGATTTATGGAGACAAATGGCTAAA 422
Qy 413 ATTTGGATACTTGAATTTCTGAGTGCCTCAAAAGTTTCAATCTCCAGGTTTAAATCCGAGAG 472
Db 423 ATATGTGCTTGGAAAGTGTCTAGTCCCAAGAAATTTGGACATATAGCTCTTATAGGCGC 482
Qy 473 GAAAGAAATGGAGGATGCCATCACATTCCTCCGTTCCGAAAGCCGGATCTCCGGTCAATATT 532
Db 483 GATGAAGTTCTTCGCTCTTAAATTTTATCCGCTCATCTCTCTGGTGAGCCTGTATAATT 542
Qy 533 ACAAGATCATTTATGGCATTTATAATTTTCGATCATGATAAGAACATCCGTTGGTAA---T 589
Db 543 ACGGAAAGGATCTTTTGTTCACAGCTCCATGACATGTAGATCAGCGTTTGGGCAAGTA 602
Qy 590 TGTAAGCAAAAGAAAGATGCTGAGTGTGGCGATGCTCAATGAGGACGACGAGCT 649
Db 603 TTCAAGGAGCAAGACAAATTTATCAACTAATTAAGAAGTTTATCTCTTTAGCAGAGGG 662
Qy 650 TTTGACCCGCGAGCGCTTTTCCGACGTGGAATTTACTTCACTATATCATTTGGAGCTGAG 709
Db 663 TTTGATGTGGCTGACATATTCCTTTCATCAAGTCTCTTCACTGCTCAGTGAATGAAG 722
Qy 710 TCAAAACCCAGGCGTTTTCGATCAGGAGATTGACGATATATCTTGAAGAGATTTCTTAAATGA 769
Db 723 GGTAAGATTATGAATGCACACCATAAAGGTAGATGCTATTGTTGAGAATGTCTATCAACAG 782
```

```
QY 770 CAC-----AAAGCCATAAGCCTTTTGAAGCGGATAACTTAATG 808
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 783 CACAGAGAAAAATCTTGCAATTCGGAAAACTAATGGAGCGTTAGGAGGTGAAGATTAAAT 842
QY 809 GATGTTCTATTGAATCTTCARAAAAATGGAACGTTCCAGTGCAGTGACAAACGAAAGC 868
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 843 GATGTTCTTCTAAAACCTTATGAATGATGGAGCGCTTCAATTTCTTATCACCAACGCAAC 902
QY 869 ATCAAGACATCGTTTTCGAAATGTTTACTGCGGGAGCGAAACAACTTCGAAAGCTACA 928
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 903 ATCAAGCTATAATCTTTGACATGTTTCTGCTGGACAGAGACTTCATCGTCAACAATT 962
QY 929 GAATGGTAAATGGCAGAGCTGATGAAAAATCCAATCAACTGAAGTAAAGAACACAGAGAA 988
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 963 GTGTGGCTATGGTGAATGCGAAAAATCCAATCTGATTTTGCAGAAAGCTCAAGCAGAA 1022
QY 989 GTTACAGAGTATTTGGTGAATGCGAAATGCGAAAGTTGATGATCAAGATTTCATGATTTGAAA 1048
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1023 GTAAGAGATGCAATTTAGAGAAAAGAAACTTTTGTATGAAATGATGTGGAGGAGCTAAAC 1082
QY 1049 TTCTTCAAGTTAGTGTAAAGAACTCTAAGATTACATCTCCCGTT---GTCTTGATT 1105
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1083 TATCTAAAGTTAGTCATTAAGAAACTCTAAGACTTCATCCACCGTTCACACTTTTGCTC 1142
QY 1106 CCGAGGAGGTAGAGAAACACAGCAATGTGATGGAATGAAATTCATCCGAAACTCGA 1165
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1143 CCAAGAGAAATGTAGGAAGAGACAAATATAAACCGCTACACTATTCTCTGTAAGACCAAA 1202
QY 1166 ATTTGTTGAATGCTTTGGCGGATAGGAAGAGATCTTAATACCTTGTGCGAACCTGGAAG 1225
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1203 GTCATGTTAATTTTGGGCATTTGGGAGAGATCCAAATATTTGGATGATGACAGAACT 1262
QY 1226 TTTAAACCCAGAAAGTTTAAAGATTGTGCAATTTGATTAATAAGGACGACATTTGAACCTG 1285
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1263 TTTAAGCCAGAGAGATTTGAGCAGTCTCTAAGGATTTTGTGGTAAATAATTTGAAATAT 1322
QY 1286 GTACATTTGGTGAGGAAAAAGAAATATGCTCTGGCAATTAATTCAGCTATTACCAATTTG 1345
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1323 CTTCATTTGGTGGTGAAGGAGGATTTGTCCAGGAGATTTCTGTTGGTTAGCTAATGCT 1382
QY 1346 GAGTATGTCATTATAAATCTATTATATCAATTTTAAATTTGGAACTGGCCGATGGAATTACA 1405
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1383 TATTTGCCATTGGCTCAATTACTTTATCACTTTTGAATGGAACTCCCACTGGAATCAA 1442
QY 1406 CCTCAAAACCTTGATATAGCTGAAGCTATTGGCGGTGCTCTCAGGAAAAAAATAGATCTT 1465
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1443 CCAAGCGACTTGGACTTGACTGAGTTGGTTGGAGTAACTGCGGCTAGAAAAAGTGACCTT 1502
QY 1466 AAGTTGATTCCTATTCCATATCAA 1489
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1503 TACTTGGTTGGCACTCCCTTATCAA 1526
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 6
US-11-116-881A-162
; Sequence 162, Application US/11116881A
; Publication No. US20060041949A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Dongmei
; APPLICANT: Nieleen, Mark T.
; TITLE OF INVENTION: Nicotiana Nucleic Acid Molecules and Uses Thereof
; FILE REFERENCE: 07678/141014
; CURRENT APPLICATION NUMBER: US/11/116,881A
; CURRENT FILING DATE: 2005-04-27
; PRIOR APPLICATION NUMBER: 60/665,451
; PRIOR FILING DATE: 2005-03-24
; PRIOR APPLICATION NUMBER: 60/665,097
; PRIOR FILING DATE: 2005-03-24
; PRIOR APPLICATION NUMBER: 60/646,764
; PRIOR FILING DATE: 2005-01-25
; PRIOR APPLICATION NUMBER: 60/607,357
; PRIOR FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: 60/566,235
```

```
; PRIOR FILING DATE: 2004-04-29
; PRIOR APPLICATION NUMBER: 10/934,944
; PRIOR FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: 10/943,507
; PRIOR FILING DATE: 2004-09-17
; PRIOR APPLICATION NUMBER: 60/503,989
; PRIOR FILING DATE: 2003-09-18
; PRIOR APPLICATION NUMBER: 60/485,368
; PRIOR FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: 60/418,933
; PRIOR FILING DATE: 2002-10-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2300
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 162
; LENGTH: 1538
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
US-11-116-881A-162

Query Match      23.7%; Score 410; DB 16; Length 1538;
Best Local Similarity 58.1%; Pred. No. 4.4e-91;
Matches 816; Conservative 0; Mismatches 555; Indels 33; Gaps 4;

QY 113 CCTCAGGGCCATGGAAAGTTTCTTATCATAGTAATCTTCCTCATTTTACTACTCACTTCT 172
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 129 CCACCAGGTCCATGGAAACTACCAATACTAGGAAGTATGCTTCATATGGT-----TGGT 182
QY 173 GATCTAGGCCATGAACGTTTGTAGAGCCTTGCTCAAAATTTATGGACCTGTTTATGAGTCTT 232
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 183 GGACTACACACCATGCTTTAGAGATTTAGCCAAAATATGGACCCGCTTATGCACCTT 242
QY 233 CAAATTGGCCAAAGTTTTCAGCTGTTGTCTATTCTTCAGCTGGAAGCAGCAAGAGAGGTTATG 292
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 243 CAAATTAGGTGAAGTTTCTGCAGTTGTGGTTACTTCTCTGTATATGGCAAAAGAGAGTACTA 302
QY 293 AAAACTCAGGCTGATGCTTCGCCCAACGCCCTATCGTCTTGGACGACAGATTCGTGTTT 352
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 303 AAAACTCATGACATGCTTTTCGCGTCTAGGCTTAGCCCTTTTGGCCCCGGAGATTTGCTGT 362
QY 353 TATAATCGGAAGATGCTCTGTTTTCATATGGAGATCACTGGAGGCGAGATGAAGAAA 412
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 363 TACAATAGTCTGATCTTGCGCTTTTGCCCCCTATGGCGATTTTGGAGACAATGCGTAAA 422
QY 413 ATTTGGATATCTTGAATTTCTGAGTGCCAAAAGAGTTCAATCTCCAGGTTAATCCGAGAG 472
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 423 ATATGTGCTTTGGAAAGTCTCAGTGCCAAAGAAATGTTTCGGACATATAGCTCTATTAGGCGC 482
QY 473 GAAGAAATGGAGGATGCCATCACATTCCTCGTTTCCGAAAGCCGGATCTCCCGTCAATATT 532
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 483 GATGAAGTTCTTCGTCTCCTTAATTTTATCCGGTCACTCTTCGTGAGCGCTGTTAATATT 542
QY 533 ACAAAGATCATTTTATGGCAATTATAATTTTCGATCATGATAAGAACAATCCGTTGGTAA---T 589
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 543 ACGGAAGGATCTTTTGTTCACAAGCTCCATGACATGTAGATCAGCGTTTGGGCAAGTA 602
QY 590 TGTAAAGCAAAAGAAAGATTTGCTGAGTGTTCGCGATGCGAGTCAATGAGCAGCAGAGT 649
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 603 TTCAAGGAGCAAGACAAAATTTATACAACATAAATAAGAAAGTTATATCTCTTAGCAGGAGG 662
QY 650 TTTCGCCCGCAGAGCGCTTTTCCGACGTTGGAATTTACTTTCATATATCATTCGAGCTGAG 709
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 663 TTTGATGTGGCTGACATATTCCTTTCAACAAGTCTCTTCATGTCTCAGTGGAAATGAAG 722
QY 710 TCAAAACCCAGCGGTTTTCATCAGGAGATTTGACGATATATCTTTGAAGAGATTCCTTAATGA 769
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 723 GGTAAGATTATGAATGCACACACCATTAAGGTAGATGCTATTGTTGAGAAATGTCATCAACGAG 782
QY 770 CAC-----AAAGCCATAAGCCTTTTGAAGCGGATAACTTAATG 808
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 783 CACAGAGAAAAATCTTGCAATTTGGGAAAACTAATGGAGCGTTAGGAGGTGAAGATTAAAT 842
QY 809 GATGTTCTATTGAATCTTCAAAAAAATGGAAGAGTTCAGTCCAGTCCAGTGAACAAACGAAAGC 868
```

Db 843 GATGTTCTTCTTAAACTTATGAATGATGGAGGCTTCAATTTCTCTATCAACCAACGACAAC 902
QY 869 ATCAAGCATCCGTTTTCGAAAGTGTTCCTGCGGAGCGGAAACAACCTTCGAAGCTACA 928
Db 903 ATCAAGCTATAATCTTTGACATGTTTGGCTGCTGGAAACAGAGACTTCATCGTCAACAAT 962
QY 929 GAATGGGTAAATGGCAGAGCTGATGAAATAATCCAACTGAACTAAGAAAGCAACAAGAA 988
Db 963 GTCTGGCTATGTTGGAATGTTGAAATGTTGAAATATCCAACTGATTTGGAAGCTCAAGCAGAA 1022
QY 989 GTTAGACAAGTATTTGGTGAATGGGAAAGTGTGATGAATCAAGATTTCAATGATTTGAAA 1048
Db 1023 GTAAGAGATGCAATTTAGAGAAAAGAACTTTTGTGATGAAAATGATGGAGGAGCTAAAC 1082
QY 1049 TTCTTCAAGTTAGTGTAAAGAACTCTAAGATTTACATCTCCGGT---GTCCTGATT 1105
Db 1083 TATCTAAGTTAGTCAATTAAGAAACTCTAAGACTTCAATCCCGGTTCCACTTTTGGCTC 1142
QY 1106 CCGAGGAGGTAGAGAAACAAACAGAAATTTGATGGATATGAAATTCATCCGAACACTCGA 1165
Db 1143 CCAAGAGATGTAGGGAAGAGACAATATATAACGGCTACACTATTTCTGTAAGACAAA 1202
QY 1166 ATTTGTTGTAATCTTTGGGCGATAGGAAGAGATCTTAATATCTTTGGTGGAACTCGGAAAG 1225
Db 1203 GTCATGGTTAATGTTTGGGCAATTTGGGAAGAGATCCAAATATTTGGGATGATGAGAACT 1262
QY 1226 TTAAACCCAGAAAGTTTAAAGATTTGCAATTTGATTAAGGGAGGACATTTGAACCTG 1285
Db 1263 TTTAAGCCAGAGAGATTTGAGCAGTGTCTTAAGGATTTTGTGGTAATATTTTGAATAT 1322
QY 1286 GTACCAATTTGGTGCAGAAAAGAAATATGCTCTGGCAATTTACTTCACTATTACCAATTTG 1345
Db 1323 CTTCATTTGGTGTGGAAGGAGATTTGTCAGGAGATTTCTGTTGGTTAGCTAATGCT 1382
QY 1346 GAGTATGCTAATATAAATCTATTAATCAATTTTAAATGGGAATGCGCGATGGAATTA 1405
Db 1383 TATTTGCCATTTGCTCAATTAATTTACTTTTATCACTTTTGAATGGAACTCCCACTGGAATCAA 1442
QY 1406 CCTCAACACTTGATATGCTAGTCAAGCTATTTGGCGGTCTCTCAGGAAAAAATAGATCTT 1465
Db 1443 CCAAGCGATGAGCTGACTGATGTTGGTGGAGTAACTGCGCTAGAAAAAGTGACCTT 1502
QY 1466 AAGTTGATTCCTATTCCATATCAA 1489
Db 1503 TACTTGGTTCGACTCTTATCAA 1526

RESULT 7

US-10-686-947-209
; Sequence 209, Application US/10686947
; Publication No. US20040162420A1
; GENERAL INFORMATION:
; APPLICANT: Profigen Inc.
; TITLE OF INVENTION: Cloning of Cytochrome P450 Genes from Tobacco
; FILE REFERENCE: 79601
; CURRENT APPLICATION NUMBER: US/10/686,947
; PRIOR FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: US 10/387346
; NUMBER OF SEQ ID NOS: 298
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 209
; LENGTH: 1673
; TYPE: DNA
; ORGANISM: NICOTIANATABACUM
US-10-686-947-209

Query Match 23.1%; Score 400.4; DB 8; Length 1673;

Best Local Similarity 57.7%; Pred. No. 1.2e-88;

Matches 810; Conservative 0; Mismatches 561; Indels 33; Gaps 4;

QY 113 CCTCAGGGCCATGGAAGTTTCTCATCATAGGTAACTTCTCTCAATTTACTCACTTCT 172

Db 135 CCACCAAGTCCATGGAACTACCAATACTAGGAAGTATGCTTCATATGGT-----TGGT 188
QY 173 GATCTAGGCGATGAACGTTTGTAGAGCTTTGGCTCAAAATTTATGGACCTGTATGAGTCTT 232
Db 189 GGACTACCAACACCATGCTCCTTAGAGATTTAGCCAAAAAATATGGACCACTTATGCACTT 248
QY 233 CAAATTTGGCCAAAGTTTTCAGCTGTGTCATTTCTTCAAGCTGAAGCAGCCAAAGAGGTATG 292
Db 249 CAAATAGGTGAAGTTTCTCGGTTTGGTTACTTCTCCGTATACGGCAAAAGATATTA 308
QY 293 AAAAATCAGGCTGATGCCCTTTCGCCAACGCCCTTATCGTCTTGGACGCACAGATTTGTTTT 352
Db 309 AAAAATCATGACATCGCTTTTGGCTTAGGCCCTAGCTTTTGGCCCCGGAGATTTGCTGT 368
QY 353 TATAATCGGAAAGATGCTGTTTGGTTTCAATGAGATCACTGGAGGAGATGAAGAAA 412
Db 369 TACAATAGGTCTGATCTAGCTTTTGGCCCTATAGGCACTATTGGAGACAAATGCGTAAA 428
QY 413 ATTTGATACTTGAATTTCTGAGTGCACAAAAGTTCAATCTCCAGGTTAATCCGAGAG 472
Db 429 ATATGTTCTTGGAAAGTCTCAGTGCACAAAGATTTTGGACATTTTACTCTTATTAGCGCG 488
QY 473 GAAGAAATGGAGATGCCATCACATTCCTCCGTTTCGAAAGCCGATCTCCGGTCAATATT 532
Db 489 AATGAAGTTCTTCGTTCTCATTAATTTATCCGCTCATCTTCTGGTGAACTATTAATGTT 548
QY 533 ACAAGATCATTTATGGCAATTAATTTTCAATGATGATGAAGAACATCCGTTGGTAA---T 589
Db 549 ACGGAAAGGATCTTTTGTTCACAGCTCCATGACATGATAGATCAGCGTTTGGCAAGTG 608
QY 590 TGTAGCAAAAAGAAAGATTTGCTGAGTGTTCGCGATGCGATCAATGAGGCGACGAGT 649
Db 609 TTCAAAGACAAAGACAAATTTATACAACTAAATAAGAAAGTGATATCTTTAGCAGGAGG 668
QY 650 TTTGGCACGCGACAGCGCTTTTCCGACGTGGAAATTACTTCACTATATCATTTGAGCTGAG 709
Db 669 TTTGATGTGCTGACATATTCCTTCACTGAAGTTTCTTCACTGCTCAGTGAATGAAG 728
QY 710 TCAAAACCCAGCGGTTTGCATCAGGAGATGACGATATACCTTTGAAGAGATTTCTTAATGA 769
Db 729 GGTAAAGATTATGAATGCACACCATAAAGTAGATGCCATTTGTGGAGAAATGTCATCAATG 788
QY 770 CAC-----AAAGCCATAAGCCTTTTGAAGCGGATTAACCTTAATG 808
Db 789 CACAAGAAAAATCTTGCATTTGGGAAAACTAATGGAGCGTTTAGGAGGTGAAGATTTAATTT 848
QY 809 GATGTTCTTATGAATCTTTCAAAAAATGGAACGTTTCCAGTGCCAGTGAACAAACGAAAGC 868
Db 849 GATGTTCTTAAAGACTTATGAATGATGGAGCGCTTCAATTTCTTATCACCAACGACAAC 908
QY 869 ATCAAAGCATCCGTTTTCGAAATGTTTACTGCGGAGCGGAAACAACTTTCGAAAGCTACA 928
Db 909 ATCAAAGCTATAAATTTTGGACATGTTTGTGCGCGGACAGAGACTTCACTCGTCAACAAAT 968
QY 929 GAATGGGTAAATGGCAGAGCTGATGAAAAATCCAACTGAACCTAAGAAAAAGCAACAAGAA 988
Db 969 GTTGGGCTATGTTAGAAATGGTGAANAATCCAGCCGTTATTCGCAAGAGCTCAAGCAGAA 1028
QY 989 GTTAGACAAGTATTTGGTGAATGGAAGTGTGATGAATCAAGATTTCTCATGTTTGAATA 1048
Db 1029 GTAAGAGAGCATTTTAGAGAAAAGAAACTTTTCGATGAATAATGATGTGGAGAGCTAAAC 1088
QY 1049 TTTCTTCAAGTTAGTGGTTAAAGAAACTCTAAGATTTACATCTCCGTT---GTCCTGATT 1105
Db 1089 TACCTAAAAGTTAGTAATAAAGAAAACTCTAAGACTTCTATCCACCGGTTCCACTTTTGCTC 1148
QY 1106 CCGAGGAGTGTAGAGAAACACACGAATTTGATGATATGAATTAATTCATCCGAACTCGA 1165
Db 1149 CCAAGAGAATGTAGGGAAGAGACAAATATAAACGGCTACACTATTTCTCTGTAAAGACCAA 1208
QY 1166 ATTTGTTGTAATGCTTTGGCGGATAGGAAGATCTTAATATCTTTGTCGGAACCTGGAAG 1225

Db 1209 GTCATGGTTAAATGTTTGGGCTTTGGGAGAGATCCAAATATTTGGAATGACCGAGAACT 1268
Qy 1226 TTTAAACCCAGAAAGGTTTAAAGATTGTGCAATTGATTAATAAGGACGACATTTGAACTG 1285
Db 1269 TTTATGCCAGAGAGATTTGAGCAGTGCTTAAGGATTTTGGTAAATAATTTGAATAT 1328
Qy 1286 GTACATTTGGTGCAGGAAAAAGAAATATGCTCTGGCAATTACTTACGCTATTACCAATTTG 1345
Db 1329 CTTCCATTTGGTGGCGGAGGAGATTTGCTCTGGGATTTCTGTTGGCTTAGCTAATGCT 1388
Qy 1346 GAGTATGTCATTATAAATCTATTATATCAATTTTAAATTGGGAACGCGCATGSAATTACA 1405
Db 1389 TATTTGCCATTGGCTCAATTACTATATCACTTCGATTGGAAATCCCTGCTGGAATCGAA 1448
Qy 1406 CCTCAAAACATTTGATATGATGAAGCTATTGGCGGTGCTCTCAGGAAAAAATATAGATCTT 1465
Db 1449 CCAAGCGACTTGGACTTGACTGAGTTGGTGGAGTAACTGCCGCTAGAAATAAGTGACCTT 1508
Qy 1466 AAGTTGATTCCTATTCCATATCAA 1489
Db 1509 TACTTGGTTGCGACTCCTTATCAA 1532

RESULT 8

US-10-686-947-211

; Sequence 211, Application US/10686947
; Publication No. US20040162420A1
; GENERAL INFORMATION:

; APPLICANT: Profigen Inc.
; TITLE OF INVENTION: Cloning of Cytochrome P450 Genes from Tobacco

; FILE REFERENCE: 79601
; CURRENT APPLICATION NUMBER: US/10/686,947
; PRIORITY FILING DATE: 2003-10-16

; PRIOR APPLICATION NUMBER: US 10/387346
; PRIOR FILING DATE: 2003-03-12

; NUMBER OF SEQ ID NOS: 298

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 211

; LENGTH: 1673

; TYPE: DNA

; ORGANISM: NICOTIANATABACUM

US-10-686-947-211

Query Match 23.1%; Score 400.4; DB 8; Length 1673;

Best Local Similarity 57.7%; Pred. No. 1.2e-88;

Matches 810; Conservative 0; Mismatches 561; Indels 33; Gaps 4;

Qy 113 CTTCCAGGCCCATGGAAGTTTCTATCATAGTAAATCTTCTCATTTTATTTACTCACTTCT 172
Db 135 CCACAGGTCATGGAACTACCAATACTAGGAATGCTTCATATGTT-----TGGT 188
Qy 173 GATCTAGGCCATGAACGTTTATAGAGCTTTGGCTCAAAATTTATGGACCTGTTATAGTCTT 232
Db 189 GGACTACACACCATGCTCTTAGAGATTTAGCCAAAAATAATGGACCACTTATGCACCTT 248
Qy 233 CAAATGGCCAAAGTTTCAGCTGTTGTCATTTCTTACGCTGAAGCAGCCAAAGAGTTATG 292
Db 249 CAATTAGGTGAAGTTTCTCGGTTGGTTACTTCTCTGATACCGCAAAAGAGTATTA 308
Qy 293 AAAACTCAGGCTGATGCTCTGCCCAACGCCCTATGCTCTGGAGCGACAGATTGTTGTT 352
Db 309 AAATCATGACATCGCTTTGGCTTAGGCTTAGGCTTTGGCCCGGAGATTGCTGT 368
Qy 353 TATAATCGGAAAGATGCTGTTGTTGCTTCATATGGAGATCACTGGAGGAGATGAAGAAA 412
Db 369 TACNATAGTCTGATCTAGCTCTTTGCCCTATGGCGACTATTGGAGACAAATGCGTAAA 428
Qy 413 ATTTGATACTTGAATTTCTGAGTGCAGGAAAGTTCAATCTTCAGGTTAATCCGAGAG 472
Db 429 ATATGTGCTTGGAGTGTCTAGTGCAGAAAGATGTTCCGACATTTAGCTCTATTAGGCGG 488
Qy 473 GAAGAAATGGAGATGCCATCACTTCTCCGTTCCAAAGCCGGATCTCCGGTCAATATT 532

RESULT 9

US-10-686-947-231

Db 489 AATGAAGTCTCTGCTCATTAATTTTATCCGGTCATCTTCTGGTGAACCTATTAAATGTT 548
Qy 533 ACAAGATCAITTTATGGCAITTAATAATTTTCGATCATGATGAAGAACATCCGTTGGTAA--T 589
Db 549 ACGAAAGGATCTTTTGTTCACAAGCTCCATGACATGTAGATCAGCGTTTGGGCAAGTG 608
Qy 590 TGTAAAGCAAAAGAAAGATTTGCTGAGTGTTCGCCGATGCAATGAGGCGAGCAGGAGT 649
Db 609 TTCAAAGAGCAAGCAAAATTTATACAATAATTAAGAAGTGATACTCTTAGCAGAGGG 668
Qy 650 TTTGSCACCGCAGACGCTTTTCCGACGCGGAAATTTACTTCACTATATCATTTGGAGCTGAG 709
Db 669 TTTGATGTTGGCTGACATATTCCCTTCACTGAAGTTTCTTCATGTGCTCAGTGAATGAAG 728
Qy 710 TCAAAACCCAGCGGTTTGCATCAGGAGATTAACGATATATCTTGAAGAGATTCTTTAATGAA 769
Db 729 GGTAAAGATTATGAATGCACACCATAAAGGTAGATGCCATTTGTTGAGAATGTTCATCAATGAG 788
Qy 770 CAC-----AAAGCCCAATAGCCCTTTTGAAGCGGATAACTTAATG 808
Db 789 CACAAGAAAAATCTTTGCAATTTGGGAAAACTAATGGAGCGTTAGGAGGTGAAGATTTAAT 848
Qy 809 GATGTTCTATTGAATCTTCAAAAAAATGAAACGTTCCAGTGCCAGTGACAAAAACGAAAGC 868
Db 849 GATGTTCTTAAAGACTTATGAATGATGGAGGCTTCAATTTCTCTATCACCAACGACAAC 908
Qy 869 ATCAAGCATCCGTTTTCAAAATGTTTACTGCGCGGAGCGAAACAACTTTCGAAAGCTACA 928
Db 909 ATCAAGCATAAATTTTTCGATGTTTCTGCTCGGAGACAGACTTTCATCGTCAACAAAT 968
Qy 929 GAATGGGTAAATGGCAGAGCTGATGAAAAATCCAACTGAATTAAGAAAAAGCACAAGAAA 988
Db 969 GTGTGGGCTATGGTAGAAAAATCGAGCCGTTATTCGCGAAAGCTCAACAGAA 1028
Qy 989 GTTAGACAAGTATTGTTGGAATGGAAAAAGTTGATGAATCAAGATTTTCATGATTGAAA 1048
Db 1029 GTAGAGAGCATTTAGAGGAAAGAAACTTTCGATGAAATATGATGTGGAGGAGTAAAC 1088
Qy 1049 TCTTCAAGTTAGTGGTTAAAGAACTCTAAGATTATACCTCTCCGGTT---GCTTGTATT 1105
Db 1089 TACCTAAAGTTAGTAATAAAGAAACTCTAAGACTTTCACCCGTTTCCACTTTTGCTC 1148
Qy 1106 CCGAGGAGTGTAGAGAAACAACGATTCATGATATGAAATTCATCCGAAACACTCGA 1165
Db 1149 CCAAGAGATGTAGGGAAGAGACAAATATAACCGCTACACTATTCCTGTAAGACCAAA 1208
Qy 1166 ATTGTTGTGAATGCTTTGGCGGATAGGAAAGAGATCCTAATACCTTGGTCGGAACCTGGAAG 1225
Db 1209 GTCATGGTTAATGTTTGGGCTTTGGGAAGAGATCCAAATATTTGGAATGACGAGAACT 1268
Qy 1226 TTTAAACCAGAAAGGTTTAAAGATTGTCGAATTTGATTAATAAGGAGCGACATTTGAACTG 1285
Db 1269 TTTATGCCAGAGAGATTTGACGAGTGCTCTAAGGATTTTGTGTTGTAATAATTTTGAATAT 1328
Qy 1286 GTACCATTTGTCAGGAAAAAGAAATATGCTCTGCAATTTACTTTCAGCTATTACCAATTTG 1345
Db 1329 CTTCCATTTGGTGGCGAAGGAGGATTTGCTCTGGGATTTGCTTTGGCTTAGCTAATGCT 1388
Qy 1346 GAGTATGTCATTATAAATCTATTATATCATTTTAAATCGGAACTGGCCGATGGAATTAACA 1405
Db 1389 TATTTGCCATTTGGCTCAATTTACTATATCATTCGATTGGAAACTCCCTGCTGGAATCGAA 1448
Qy 1406 CCTCAACACTTGATATGACTGAAGCTATTGGCGGTGCTCTCAGGAAAAAATATAGATCTT 1465
Db 1449 CCAAGCGACTTGGAGCTTGAAGTGTGTTGGAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1508
Qy 1466 AAGTTGATTCCTATTTCATATCAA 1489
Db 1509 TACTTGGTTCGACTCCTTATCAA 1532

```
; Sequence 231, Application US/10686947
; Publication No. US20040162420A1
; GENERAL INFORMATION:
; APPLICANT: Profigen Inc.
; TITLE OF INVENTION: Cloning of Cytochrome P450 Genes from Tobacco
; FILE REFERENCE: 79601
; CURRENT APPLICATION NUMBER: US/10/686,947
; CURRENT FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: US 10/387346
; PRIOR FILING DATE: 2003-03-12
; NUMBER OF SEQ ID NOS: 298
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 231
; LENGTH: 1673
; TYPE: DNA
; ORGANISM: NICOTIANATABACUM
US-10-686-947-231

Query Match      23.1%; Score 400.4; DB 8; Length 1673;
Best Local Similarity 57.7%; Pred. No. 1.2e-88;
Matches 810; Conservative 0; Mismatches 561; Indels 33; Gaps 4;

QY 113 CCTCAGGGCCATGGAAGTTTCTATCATAGGTAAATCTTCTCATTTTACTCACTTCT 172
DB 135 CCACCAAGTCCATGGAACCTACCAATACCTAGGAAGTATGCTTCATATGGT-----TGGT 188

QY 173 GATCTAGGCCATCAACGTTTTAGAGCCTTGGCTCAAAATTTATGGACCTGTATGACTCTT 232
DB 189 GGNCTACCACCATGCTCTTAGAGATTTTAGCCAAAATATATGGACCTTATGCACTT 248

QY 233 CAAATTTGGCCAAAGTTTTCAGCTGTGTGTCATTTCTCAGCTGAAGCAGCCAAAGAGGTATG 292
DB 249 CAATTAGGTGAAGTTTCTGCGGTGTGGTTACTTCTCCTGATACGGCAAAAGATTTA 308

QY 293 AAAACTCAGCTGATGCCCTTGGCCAAAGCCCTATCGTCTTGACCGCAGACAGATGTGTTT 352
DB 309 AAAACTCATGACATCGCTTTTGGCTAGGCCCTAGCCCTTTTGGCCCGGAGATTTGCTGT 368

QY 353 TATAATCGGAAGATGCTTTGTTGCTTCATATGGAGATCACTGGAGGAGATGAAGAA 412
DB 369 TACAATAGTCTGATCTAGCCTTTGCCCCATATGGCGACTATTTGGAGACAAATGCGTAAA 428

QY 413 ATTTGGATCTTGAATTTCTGAGTGCACAAAAGTTCAATCTCTCAGGTTTAATCCGAGAG 472
DB 429 ATATGTGCTTGAAGTGTCTAGTGCCCAAGATGTTTCGGACATTTAGCTCTATTAGGCGG 488

QY 473 GAAGAAATCGAGATGCCATCACATTCCTCGGTTGAAAGCCGGATCTCCGGTCAATATT 532
DB 489 AATGAAGTTCTTCGCTCATTAATTTTATCCGGTCACTCTCTGGTGAACCTATTAAATGTT 548

QY 533 ACAAGATCATTTATGGCATTATAATTTTCGATCATGATAAGAACATCCGTTGGTAA---T 589
DB 549 ACGGAAGGATCTTTTGTTCACAGCTCCATGACATGTAGATCAGCGTTGGGCAAGTG 608

QY 590 TGAAGCAAAAAGAAAGATTGCTGAGTGTGCGGATGCGAGTCAATGAGGCGAGCGAGT 649
DB 609 TTCAAAGAGCAAGACAAATTTATACAACTAATTAAGAAAGTGATCTCTTAGCAGGAGG 668

QY 650 TTGGCACCAGCAGACGCTTTTCGAGCTGGAATTAATCTCACTATATCATTTGAGCTGAG 709
DB 669 TTGTGATGTGGCTGACATATTCCTTCCTCACTGAAGTTTCTTCATGTGCTCAGTGGAAATGAAG 728

QY 710 TCAAAACCCAGGCTTTTGATCAGGAGATTGAGATATATCTTGAAGAGATTTCTTAATGAA 769
DB 729 GGTAAGATTATGAATGCACACCATAGGAGTAGATGCCATTTGTTGAGAAATGTCATCAATGAG 788

QY 770 CAC-----AAAGCCAAATAAGCCTTTTGAAGCGGATAACTTTAATG 808
DB 789 CACAAGAAAAATCTTCGAATTTGGGAAAACCTAATGGAGCGTTAGGAGTGAAGATTTAAT 848

QY 809 GATGTTCTAATGAATCTTCAAAAAATGAAACGTTTCCAGTGCCAGTGAACAAACGAAAGC 868
DB 849 GATGTTCTTAAAGACTTATGAATGATGAGGCGCTTCAATTTCTTATCATCAACCAACGACAA 908
```

```
QY 869 ATCAAGCATCCGTTTGGCAATGTTTACTGCGGGGAGCGAACAACACTTCGAAGCTACA 928
DB 909 ATCAAGCCCATTAATTTTGGCATGTTTGTGCGGGGACAGAGACTTCATCGTCAACAAT 968

QY 929 GAATGGTAAATGGCAGAGCTGATGAAAAATCCAACTGAACCTAAGAAAAAGCACAAGAGAA 988
DB 969 GTGTGGCTATGCTAGAAATGGTGAANAATCCAGCGGTATTCGGAAGCTCAAGCAGAA 1028

QY 989 GTTAGACAAGTATTTGCTGAATGGGAAAAGTTGATGAATCAAGATTTCAATGATTTGAAA 1048
DB 1029 GTAAGAGAAAGCATTTTAGAGGAAAAGAACTTTTCGATGAAAAATGATGCGAGGAGCTAAAC 1088

QY 1049 TTCTTCAAGTTAGTGGTTAAAGAAACTCTAAGATTTACATCTCTCCGGT---GTCCTGATT 1105
DB 1089 TACCTAAAGTTAGTAATAAAGAACTCTAAGACTTCATCCACCGGTTCCACTTTTGCTC 1148

QY 1106 CCGAGGAGCTGTAGAGAAACAAACAGAAATTTGATGGATATGAAATTCATCCGAACACTCGA 1165
DB 1149 CCAAGAGAAATGTAGGGAAGAGACAAATATAAAGCGCTACACTATTCTGTAAAGACCAA 1208

QY 1166 ATGTTGTGAATGCTTTGGCGGATAGGAAGAGATCCTAATCTTTGTCGGAACCTGGAAAG 1225
DB 1209 GTCATGGTTAATGTTTGGGCTTTGGGAGAGATCCAAAATATTGGAATGACGCGAGAACT 1268

QY 1226 TTTAACCCAGAAAGGTTTAAAGATTTGCAATTTGATTATAAAGGAGCAGCATTTGAAC 1285
DB 1269 TTTATGCCAGAGAGATTTGAGCAGTCTCTAAGGATTTTGTGGTAATAATTTTGAATAT 1328

QY 1286 GTACCATTTGGTCAGGAAAAAGAAATATGCTCTGCGCATTTACTTCAGCTATTACCAATTTG 1345
DB 1329 CTTCCATTTGGTGGCGAAGGAGGATTTGCTCTGGGATTTGTTTGGCTTAGCTAATGCT 1388

QY 1346 GAGTATGCTAATTAATAATCTTATATATCAATTTTAAATTTGGAACCTGCGCGATGGAATVACA 1405
DB 1389 TATTTGCCATTTGGCTCAATTTACTATATCACTTTCGATTGGAAAATCTCCCTGCTGGAATCGAA 1448

QY 1406 CCTCAAAACACTTGATATGATCAAGCTATTGGCGGTGCTCTCAGGAAAAAATAGATCTT 1465
DB 1449 CCAAGGACTTGGACTTGACTGAGTTGGTGGAGTAACATGCGCGTAGAAAAAGTGACCTT 1508

QY 1466 AAGTTGATTCTTATCCATATCAA 1489
DB 1509 TACTTGGTTGGACTCTTATCAA 1532

RESULT 10
US-10-943-507-206
; Sequence 206, Application US/10943507
; Publication No. US2005013244A1
; GENERAL INFORMATION:
; APPLICANT: U.S. Smokeless Tobacco Company
; TITLE OF INVENTION: 83492-7270 CIP - Cloning of Cytochrome P450 Genes From Nicotiana
; FILE REFERENCE: 83492-7270
; CURRENT APPLICATION NUMBER: US/10/943,507
; CURRENT FILING DATE: 2004-09-17
; PRIOR APPLICATION NUMBER: 10/686,947
; PRIOR FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: 10/387,346
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: 10/340,861
; PRIOR FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 60/363,684
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/347,444
; PRIOR FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 60/337,684
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 206
; LENGTH: 1673
; TYPE: DNA
```

```
; ORGANISM: NICOTIANATABACUM
US-10-943-507-206

Query Match      23.1%; Score 400.4; DB 10; Length 1673;
Best Local Similarity 57.7%; Pred. No. 1.2e-88;
Matches 810; Conservative 0; Mismatches 561; Indels 33; Gaps 4;

QY 113 CCTCAGGGCCATGGAAGTTTCTATCATAGTAATCTTCTCATATTTACTCACTTCT 172
    |||||
Db 135 CCACAGAGTCTAGGAATCACTAATCACTAGGAAGTATGCTTCATATGCT-----TGGT 188

QY 173 GATCTAGGCCATGAACGTTTGTAGAGCTTGGCTCAAAATTATGGACCTGTTATGAGTCT 232
    |||||
Db 189 GGACTACCAACACCTATGCTTAGAGATTTAGCCAAAATAATATGGACCTTATGACCTT 248

QY 233 CAAATTGGCCAAAGTTTTCAGCTGTGTCTTCTTCTGAGCTGAAGCAGCAAGAGGTTATG 292
    |||||
Db 249 CAAATTAGGTGAAGTTTCTGCGGTTTGTGTTACTTCTCTGATACGGCAAAAGAGTATTA 308

QY 293 AAAAATCAGGCTGATGCCCTTCCCAACGCCCTATCGTCTTGGACGCAAGATTTGTGTTT 352
    |||||
Db 309 AAAAATCAGTACATGCTTTTGGCTTAGGCCCTAGCCCTTTTGGCCCGGAGATTTGCTGT 368

QY 353 TATATCGGAAGATGCTTGTGTTTCTTCTATGAGATCATCTGGAGCAGATGAAGAA 412
    |||||
Db 369 TACAAATAGGCTGATCTAGCCCTTTTGGCCCTTAGGGACTATTGGAGCAAAATGCGTAAA 428

QY 413 ATTTGGATCTTGAATTTCTGAGTGCACAAAAGTTCAATCTCCAGGTTAATCCGAGAG 472
    |||||
Db 429 ATATGTGCTTGGAGTGTCTAGTGCCAAAGATTTTGGACATTTAGCTCTATTAGGCGG 488

QY 473 GAAGAAATGAGGATGCCATCACTTCTCCGTTTGAAGCCGATCTCCGGTCAATATT 532
    |||||
Db 489 AATGAAGTTCTTCCGCTCATTAATTTATCCGCTCATCTTCTGCTGCAACCTATTAAATGTT 548

QY 533 ACAAGATCAATTTATGGCATTATATTCGATCATGATGAAGCAATCCGTTGGTAA---T 589
    |||||
Db 549 ACGGAAGGATCTTTTGTTCACAGCTCCATGACATGTAGATCAGCGTTTGGGCAAGTG 608

QY 590 TGTAGCAAAAGAAAGATTGCTGAGTGTGCGATGTCAGTCAATGAGCAGCAGCAGT 649
    |||||
Db 609 TTCAAGAGCAGACAAATTTATACAACTAATTAAGAAGTGTACTCTTAGCAGGAGG 668

QY 650 TTGGCAGCGCAGAGCTTTTCCGAGCTGGAATTTACTTCACTATATCATTTGGAGCTGAG 709
    |||||
Db 669 TTTGATGTGCTGACATATTCCTTCACTGAAGTTTCTTCACTGTGCTCAGTGGAAATGAAG 728

QY 710 TCAAAACCCAGGCGTTGTCATCAGGAGTTGACGATATATCTTGAAGAGATTTCTTAATGAA 769
    |||||
Db 729 GGTAAAGATTATGAATGCAACCATAGGTAGATGCCAATGTTGAGAAATGTCATCATGAG 788

QY 770 CAC-----AAAGCCAAATAGCCCTTTTGAAGCGGATTAATTAATG 808
    |||||
Db 789 CACAGAAAAATCTTCCAATTTGGAAAACTAATGGGCGTTAGGAGGTGAAGATTTAAT 848

QY 809 GATGTTCTATGAACTTCCAAAAAATGGAACCGTTTCCAGTCCAGTGCACAAACGAAAGC 868
    |||||
Db 849 GATGTTCTTAAAGACTTATGAATGATGAGGAGGCTTCAATTTCTTATCACCACGACAC 908

QY 869 ATCAAGCATCGTTTGGCAATGTTTACTGCGGGGCGCAACAACTTCGAAAGCTACA 928
    |||||
Db 909 ATCAAGCTATTAATTTTTCACATGTTTGTGCGGGGACAGAGACTTTCATCGTCAACAAT 968

QY 929 GAATGGGTAAATGGCAGAGCTGATGAAAAATCCAACTGAAGTAAAGAAAAACACAAGAA 988
    |||||
Db 969 GTGTGGCTATGTTAGAAATGGTGAATAATCCAGCGTATTCCGGAAGCTCAAGCAGAA 1028

QY 989 GTTAGCAAGATTTTGGTGAATGGGAAAGTTGATGAATCAAGATTTTCATGATTGAAA 1048
    |||||
Db 1029 GTAAGAGAGAGCTTTTAGAGGAAAGAAACTTTTCGATGAAAAATGATGTGGAGGAGCTAAAC 1088

QY 1049 TTCTTCAAGTTAGTGTATAAGAAACTCTAAGATTACATCTCCCGGTT---GTCCTGATT 1105
    |||||
```

```
Db 1089 TACCTAAAGTTAGTAATAAAGAAACTCTAAGACTTTCATCCACCGTCCACCTTTTGCTC 1148
QY 1106 CCGAGGGAGTCTAGAGAAACACACGAATTCATGATATGAAATTCATCCGAAACACTCGA 1165
    |||||
Db 1149 CCAAGAGAATCTAGGGAAGAGACAAATATAAAACGCTACACTATTCTCTGTAAGAGCAAA 1208
    |||||
QY 1166 ATTGTTGTAATGCTTGGCGATAGGAGAGATCTTATATCTTGGTCCGAAACCTCGAAAG 1225
    |||||
Db 1209 GTCATGGTTAATGTTTGGGCTTTGGGAAGAGATCCAAAATATTGGAATGAGCGCAGAACT 1268
    |||||
QY 1286 TTAAACCCAGAAAGGTTTAAAGATTGTGCAATTGATTATAAGGACGACACATTTGCAACTG 1285
    |||||
Db 1269 TTTATGCCAGAGAGATTTGACAGTGTCTTAAGATTTTGTGTTAAATTTTGAATAT 1328
    |||||
QY 1286 GTACCATTTGGTCAGGAAAAAGAAATATGTCCTGSCATTTACTTACGCTATTACCAATTTG 1345
    |||||
Db 1329 CTTCCATTTGGTGGCGAAGGAGGATTGTCCTGGGATTTGCTTTGGCTTAGCTAATGCT 1388
    |||||
QY 1346 GAGTATGTCATTAATAATCTATTATATCATTTTAAATTTGGGAACCTGCCGATGGAATACA 1405
    |||||
Db 1389 TATTGGCAATTTGGCTCAATTTACTATATCATCTTCGATTGGAAACTCCCTGCTGGAATCGAA 1448
    |||||
QY 1406 CCTCAAAACACTTGTATATGACTGAAGCTATTGGCGGTGCTCTCAGGAAAAAATATAGATCTT 1465
    |||||
Db 1449 CCAAGCGACTTTGGACTTGGACTGAGTTGGTTGGAGTTAACTGCGCTAGAAAAAGTGCACCTT 1508
    |||||
QY 1466 AAGTTGATTTCCATTTCATATCAAA 1489
    |||||
Db 1509 TACTTGGTTGGGACTCTCTATCAA 1532
    |||||
```

RESULT 11

```
US-10-943-507-208
; Sequence 208, Application US/10943507
; Publication No. US20050132444A1
; GENERAL INFORMATION:
; APPLICANT: U. S. Smokeless Tobacco Company
; TITLE OF INVENTION: 83492-7270 CIP - Cloning of Cytochrome P450 Genes From Nicotiana
; FILE REFERENCE: 83492-7270
; CURRENT APPLICATION NUMBER: US/10/943.507
; CURRENT FILING DATE: 2004-09-17
; PRIOR APPLICATION NUMBER: 10/686,947
; PRIOR FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: 10/387,346
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: 10/340,861
; PRIOR FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 60/363,684
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/347,444
; PRIOR FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 60/337,684
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 208
; LENGTH: 1673
; TYPE: DNA
; ORGANISM: NICOTIANATABACUM
US-10-943-507-208
```

```
Query Match      23.1%; Score 400.4; DB 10; Length 1673;
Best Local Similarity 57.7%; Pred. No. 1.2e-88;
Matches 810; Conservative 0; Mismatches 561; Indels 33; Gaps 4;

QY 113 CCTCAGGGCCATGGAAGTTTCTTATCATAGTAATCTTCTCATATTTACTCACTTCT 172
    |||||
Db 135 CCACAGAGTCTAGGAATCACTAATCACTAGGAAGTATGCTTCATATGCT-----TGGT 188

QY 173 GATCTAGGCCATGAACGTTTGTAGAGCTTGGCTCAAAATTATGGACCTGTTATGAGTCT 232
    |||||
Db 189 GGACTACCAACACCTATGCTTAGAGATTTAGCCAAAATAATATGGACCTTATGACCTT 248
    |||||
```


QY 233 CAAATTGGCCAAAGTTTCAGCTGTGTGTCATTTCTTTCAGCTGAAGCAGCCAAAGAGGTTATG 292
Db 249 CAATTAGGTGAAGTTTTCGGGTGTGGTTTACTTCTCTGTAGATACGCGCAAAAGATATTA 308
QY 293 AAAACTCAGCTGATGCTTCGCCAAGCCCTATCGTCTTGGACCCACAGATTGTGTTT 352
Db 309 AAAACTCATGACATCGCTTTTGGCTTAGCCCTTTTGGCCCCGGAGATTGTCTGT 368
QY 353 TATAATCGGAAAGATCTCTGTGTGTCTCATATGGAGATCAGTGGAGGCGAGATGAAGAAA 412
Db 369 TACAATAGTCTGATCTAGCTTTTGGCCCTATGGCGACTATTTGGAGACAAATGCGTAAA 428
QY 413 ATTTGATACTGAAATTTCTGATGCGCAAAAAAGTTCAATCTCTCCAGGTTAATCCGAGAG 472
Db 429 ATATGTGTCTTGAAGTGTCTAGTGCAGAGTGCAGAAAGTTTGGGACATTTAGCTCTATTAGCGG 488
QY 473 GAAGAAATCGAGATGCCATCAATTCCTCCGTTTCAAAGCCGGATCTCCGGTCAATATTT 532
Db 489 AATGAAGTTCTTTCGTCTCAATTAATTTATCCGGTCACTCTTCTGGTGAACCTATTATGTT 548
QY 533 ACAAGATCATTTATGGCATTATAATTTTCGATCATGATAAGAACATCCGTTGGTAA--T 589
Db 549 AGGAAAGATCTTTTGTTCACAGCTCATGACATGTAGATCAGCGTTTGGCGAAGTG 608
QY 590 TGTAGCAAAAAAGATTTGCTAGTGTTCGGATGCGAGTCAATGAGGCGAGCGAGT 649
Db 609 TTTCAAAGAGCAAGCAAAATTTATACAACTAAATTAAGAAGTGATCTCTTAGCAGGAGG 668
QY 650 TTTGGCACCGCAGACGCTTTTCGAGCGTGGAAATTTACTTCACTATATCAATTGGAGCTGAG 709
Db 669 TTTGATGTGGCTGACATATTTCCCTTCACTGAAGTTTCTTCATGTGTCTAGTGGAAATGAAG 728
QY 710 TCAAAACCCAGCGGTTTGCATCAGGAGATTGACGATATCTTGAAGAGATTTCTTAATGAA 769
Db 729 GGTAAAGTTATGAATGCACACCATAGGTAGATGCCATTTGTTGAGATGTCTATCAATGAG 788
QY 770 CAC-----AAAGCCAAATAAGCCCTTTTGAAGCGGATAACTTTAATG 808
Db 789 CACAAGAAAAATCTTGCAATTTGGGAAATCTAATGGAGCGTTAGGAGTGAAGATTTAATTT 848
QY 809 GATGTTCTAATGATTTCAAATAATGAAAGTTTCCAGTGCAGTGAACAAAGCAAGC 868
Db 849 GATGTTCTTAAGACTTATGAATGATGGAGCGCTTCAATTTCTTCACTCAACCAACGACAC 908
QY 869 ATCAAGACATCGTTTTCGAAATGTTTACTGCGGAGCGGAAACAACTTTTCGAAAGCTACA 928
Db 909 ATCAAGACCAATTTTTCACATGTTTGTGCGGAGCAGAGACTTCATCGTCAACAATTT 968
QY 929 GAATGGGTAATGGCAGAGCTGATGAAAAATCCAACTGAATGAAGAAAAAGCACAAGAGAA 988
Db 969 GTGTGGCTATGTTAGAAATGTTGAAAAATCCAGCGTATTCGCGAAAGCTCAAGCAGAA 1028
QY 989 GTTAGACAAGTATTTGGTGAATTTGGGAAAGTTGATGAATCAAGATTTCAATGATTTGAAA 1048
Db 1029 GTAAGAGAGCAATTTAGAGGAAAGAACTTTTCGATGAAATGATGTGGAGGAGCTAAAC 1088
QY 1049 TTCTTCAAGTTAGTGTGTAAGAACTCTAAGATTACATCTCCCGGTT---GCTTGTGTT 1105
Db 1089 TACCTAAAGTTAGTAAATAAGAACTCTAAGACTTCATCCACCGGTTCCACTTTTGTCTC 1148
QY 1106 CCGAGGAGTGTAGAGAAACAAACAGAAATTTGATGATATGAATTTATCCGAACACTCGA 1165
Db 1149 CCAAGAGAAATGTAGGAAGAGACAAATATAAAGCGTACACTATTTCTGTAAAGACCAA 1208
QY 1166 ATTTGTTGCAATGCTTTGGCGGATAGGAAGATTCCTAATACATTGTTGGGAACTCGGAAG 1225
Db 1209 GTCATGTTTAATGTTTGGGCTTTGGGAGAGATTCAAAAATATTGGAATGACGCAAGAACT 1268
QY 1226 TTTAAACCCAGAAAGTTTAAAGATTGTGCAATTTGATTATTAAGGAGGACGACATTTCACTG 1285
Db 1269 TTTATGCCAGAGAGATTTGAGCAGTGTCTTAAGGATTTTGTGGTAATTAATTTGAATAT 1328
QY 1286 GTACCAATTTGTGTCAGGAAAAAGAAATATGTCCTGGCATTTACTTTCAGCTATTACCAATTTG 1345

Db 1329 CTTCCATTTGGTGGCGAGGAGGATTTCCTCGGATTTTCGTTGGCTTAGCTAATGCT 1388
QY 1346 GAGTATGCTATTAATAATCTATATATCAATTTTAATTTGGAACTGGCCGATGAATATACA 1405
Db 1389 TATTTGCCATTTGGCTCAATTTACTATATCACTTCGATTGGAAACTCCCTGCTGGAATCGAA 1448
QY 1406 CTTCAAAACACTTTGATATGACTGAAGCTATTTGGCGGTGCTCTCAGGAAAAAATAGATCTT 1465
Db 1449 CCAAGGACTTGGACTTGGCTTGGTGGTGGAGTAACTGCCGCTAGAAAAAGTACCTT 1508
QY 1466 AAGTTGATTCCTATTTCATATCAAA 1489
Db 1509 TACTTGGTTGCGACTCTTATCAA 1532

RESULT 12
US-10-934-944-209
; Sequence 209, Application US/10934944
; Publication No. US20060037096A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Dongmei
; TITLE OF INVENTION: Cloning of Cytochrome P450 Genes From Nicotiana
; FILE REFERENCE: 07678/141008
; CURRENT APPLICATION NUMBER: US/10/934,944
; CURRENT FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: 10/686,947
; PRIOR FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: 60/503,989
; PRIOR FILING DATE: 2003-09-18
; PRIOR APPLICATION NUMBER: 60/485,368
; PRIOR FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: 10/387,346
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: 10/340,861
; PRIOR FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 10/293,252
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 60/418,933
; PRIOR FILING DATE: 2002-10-16
; PRIOR APPLICATION NUMBER: 60/363,684
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/347,444
; PRIOR FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 60/337,684
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 387
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 209
; TYPE: DNA
; LENGTH: 1673
; ORGANISM: Nicotiana tabacum
; US-10-934-944-209

Query Match 23.1%; Score 400.4; DB 11; Length 1673;
Best Local Similarity 57.7%; Pred No. 1.2e-88;
Matches 810; Conservative 0; Mismatches 561; Indels 33; Gaps 4;

QY 113 CCTCAGGGCCATGGAAGTTTCCTATCATAGTAACTCTCTCTCACTTTTACTCACTTCT 172
Db 135 CCACAGGTCCTATGGAACCTACCAATACTAGGAAGTATGCTTCATATGCT-----TGGT 188
QY 173 GATCTAGGCCATGACGTTTGTAGCCCTTGTGCTCAAAATTTATGGACCTGTTATGAGTCTT 232
Db 189 GGACTACCACACCATGTCTCTTAGAGATTTAGCCAAAAAATATGGACCACTTATGCACCTT 248
QY 233 CAAATTGGCCAAAGTTTTCAGCTGTTCTCATTTCTTTCAGCTGAAGCAGCCAAAGAGGTTATG 292
Db 249 CAATTAGGTGAAGTTTCTGCGGTTGTGGTTACTTCTCTGTATACGCGCAAGAGATTA 308
QY 293 AAAACTCAGGCTGATGCTTCGCCCAACGCCCTATCGCTTTGGAGCGCACAGATTGTGTTT 352
Db 309 AAAACTCATGACATCGCTTTTTCGCTTAGGCTTAGCCCTTTTGGCCCCGGAGATTGTCTGT 368

Db 429 ATATGTGCTTGTGAAGTGCTCAGTGCACGAATGTTGCGACATTTAGCTCTATTAGGCGG 488
QY 473 GAAGAAATGAGGATGCGATCACAATCTCCGTTTGAAGCCGGAATCTCGGTCAATATT 532
Db 489 AATGAAGTCTTGGTCTCATTAATTTATCCGGTCATCTTCTGGTGAACCTATTAAATGTT 548
QY 533 ACAAGATCAATTTATGGCAATTAATAATTTGCATCATGATAAGAACATCCGTTGGTAA--T 589
Db 549 ACGGAAAGGATCTTTTGTGTCAAGCTCCATGACATGTAGATCAGCGTTTGGGCAAGTG 608
QY 590 TGTAAACAAAGAAAGATGCTGAGTGTGCGCATGCGATGCGATCAATGAGGCGAGCAGT 649
Db 609 TTCAAAGAGCAAGCAAAATTTATACAACTAATTTAAAGAAAGTGATCTCTTAGCAGGAG 668
QY 650 TTTGGCACCAGCAGCGCTTTCCGAGCGTGGAAATTAATCTTCACTATATCAATTTGAGCTGAG 709
Db 669 TTTGATGTGCTGACATATTCCTTCACTGAAGTTTCTTCACTGTGCTCACTGGAATGAAG 728
QY 710 TCAAAACCCAGGCGTTTGCATCAGGAGATGACGATATATCTTGAAGAGATTTCTTAATGAA 769
Db 729 GGTAAGATTATGAATGCACACCAATAAGGTAGATGCCAATGTTGAGAAATGTCATCAATGAG 788
QY 770 CAC-----AAAGCCAAATAGCCCTTTTGAAGCGGATTAATTAATG 808
Db 789 CACAAGAAATCTTGCAATTTGGGAAATCTAATGAGCGTTAGGAGGTGAAGATTTTAAT 848
QY 809 GATGTTCTATTGAATCTTCAAAAAATGGAACGTTCCAGTGCAGTGACGACAAACGAAAGC 868
Db 849 GATGTTCTTGAAGCTTATGAATGATGAGGCGTTTCAATTTCTTATCACCAGCAAC 908
QY 869 ATCAAAGCATCCGTTTTCGAAATGTTTACTGCGGAGGCGAAACAACTTCGAAAGCTACA 928
Db 909 ATCAAAGCCATAATTTTGCATGTTTGTGCGGAGCAGAGACTTTCATCGTCAACAAT 968
QY 929 GAATGGGTAAATGCGAGAGCTGATGAATAATCCAACTGAACTAAGAAAGCAAGAAGAA 988
Db 969 GTGTGGCTATGTAAGAAATGTTGAAATAATCCAGCCGTATTCCGAAAGCTCAAGCAGAA 1028
QY 989 GTTAGACAAGTATTTGGTGAATGCGAAAGTTGATGAATCAAGATTTTCATGATTTGAA 1048
Db 1029 GTAAGGAAGCATTTAGAGAAAGAAAGAACTTTCGATGAATAATGATGTGGAGGAGCTAAC 1088
QY 1049 TTTCTTCAAGTTAGTGTAAAGAACTCTAAGATTATACCTCCCGGTT---GTCTTGATT 1105
Db 1089 TACCTAAAGTTAGTAATAAAGAACTCTAAGACTTTCATCCACCGGTTCCACTTTTGCTC 1148
QY 1106 CCGAGGAGTTAGAGAAACACACGAATGTAGGATATGAATTCATCCGAACACTCGA 1165
Db 1149 CCAAGAGAAATGTAGGGAAGAGACAAATATAAAGCGCTACACTATTCTGTAAAGACCAA 1208
QY 1166 ATTGTTGCAATGCTTGGGCGATAGGAAGATCCCTAATCTTGGTCCGAACCTGGAAG 1225
Db 1209 GTCATGTTAATGTTTGGGCTTTGGGAGAGATCCAAATAATTTGGAATGAGCAGAACT 1268
QY 1226 TTTAAACCCAGAAAGGTTTAAAGATTGTGCAATTTGATTATAAAGGAGCAGCATTTGAAC 1285
Db 1269 TTTATGCCAGAGAGATTTGAGCAGTGCTCTAAGGATTTTGTGGTAATAATTTGAATAT 1328
QY 1286 GTACCAATTTGGTGCAGGAAAAAGAAATATGTCCTGGCATTAATTCAGCTAATACCAAT 1345
Db 1329 CTTCCAATTTGGTGGCGGAGGAGGATTTGTCCTGGGATTTCTGTTGGCTTAGCTAATGCT 1388
QY 1346 GAGTATGCTATTATAATCTATTATATCAATTTTAAATTTGGAACTGCGCGATGGAATTACA 1405
Db 1389 TATTTGCCATTTGGTCAATTAATTAATCACTTCGATTTGGAATCTCCCTGCTGGAATCGAA 1448
QY 1406 CCTCAAAACATTTGATAGTGAAGCTATTGGCGGTGCTCTCAGGAAAAAATAATAGATCTT 1465
Db 1449 CCAAGCGACTTGAGCTTGACTGAGTTGGTGGAGTAACTGCCGCTAGAAAAAGTGACCTT 1508
QY 1466 AAGTTGAATCTTATTCATATCA 1489
Db 1509 TACTTGGTGGCACTCCTTATCAA 1532

RESULT 14
US-10-934-944-231
; Sequence 231, Application US/10934944
; Publication No. US20060037096A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Dongmei
; TITLE OF INVENTION: Cloning of Cytochrome P450 Genes From Nicotiana
; FILE REFERENCE: 07678/141008
; CURRENT APPLICATION NUMBER: US/10/934,944
; CURRENT FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: 10/686,947
; PRIOR FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: 60/503,989
; PRIOR FILING DATE: 2003-09-18
; PRIOR APPLICATION NUMBER: 60/485,368
; PRIOR FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: 10/387,346
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: 10/340,861
; PRIOR FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 10/293,252
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 60/418,933
; PRIOR FILING DATE: 2002-10-16
; PRIOR APPLICATION NUMBER: 60/363,684
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/347,444
; PRIOR FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 60/337,684
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 387
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 231
; LENGTH: 1673
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
US-10-934-944-231

Query Match 23.1%; Score 400.4; DB 11; Length 1673;
Best Local Similarity 57.7%; Pred. No. 1.2e-88;
Matches 810; Conservative 0; Mismatches 561; Indels 33; Gaps 4;
QY 113 CCTCAGGCGCCATGGAAGTTTCCTATCATAGTAACTCTCCCTCAATTTATTACTCACTCTT 172
Db 135 CCACAGGTCCTAGGAACTACCAATACTAGGAAGTATGCTTTCATATGGT-----TGGT 188
QY 173 GATCTAGGCGCATGAACGTTTGTAGAGCCTTTGGCTCAAATTTATGGACCTCTTATGAGTCTT 232
Db 189 GGACTACCAACCATGTCCTTAGAGATTAGCCAAATAATATGGACCCTTATGCACCTT 248
QY 233 CAAATGGCCAAAGTTTCAGCTGTTGTCAATTTCTTCAAGCTGAAGCAGCCAAAGAGGTTATG 292
Db 249 CAATTAGGTGAAGTTTCTGCGGTTGTTGTTACTTCTCTGATACGCAAAAGAAATATTA 308
QY 293 AAAACCTCAGGCTGATGCGCTTCCGCCAAAGCCCTATCGTCTCGACGACAGATTGTGTTT 352
Db 309 AAAACCTCATGACATCGCTTTTGGCTGTAGGCTAGCCCTTTTGGCCCGGAGATTGTCTGT 368
QY 353 TATAATCGGAAAGATGTCTTGTGTTGCTTATATGGAGATCACTGGAGGCGAGATGAAGAAA 412
Db 369 TACAATAGTCTGATCTAGCCTTTTGGCCCTATGCGACTATTGGAGACAAATGCGTAAA 428
QY 413 ATTTGGATATCTTGAATTTCTGAGTCCCAAAAAAGTTTCAATCTCCAGGTTAATCCGAGAG 472
Db 429 ATATGTGCTTGAAGTGCTCAGTCCCAAGAAATGTTTCGACATTTAGCTCTATTAGGCGG 488
QY 473 GAAGAAATGGAGGATGCCATCACATTCCTCCGTTTCAAGAGCCGATCTCCGGTCAATATT 532
Db 489 AATGAAGTTCTTCGCTCATTAATTTTATCCCGTCACTCTTCGGTGAACCTTATTAATGTT 548
QY 533 ACAAGATCATTTATGGCATTAATAATTTTCGATCATGATAAGAACATCCGTTGGTAA---T 589

```
Db 549 ACGGAAGGATCTTTTGTTCACAGCTCCATGACATGTAGATCAGCGTTTGGCGAAGTG 608
Qy 590 TGTAGCAAAAGAAAGATGCTGAGTGTGCGAGTCAGTCAATGAGCGAGCGAGT 649
Db 609 TTCAAGAGCAAGCAAAATTTATACAACTAATTAAGAAGTGATCTTAGCAGGAGG 668
Qy 650 TTTGCAACCGCAGACGCTTTTCCGACGTGGAAATTAATCTCACTATATCATTTGGAGCTGAG 709
Db 669 TTTGATGTGCTGACATATTCCTTCACTGAAATTTCTTCATGTCTCAGTGAATGAAG 728
Qy 710 TCAAAACCGAGCGTTTGCATCAGAGATGACGATATCTTGAAGAGATCTTAAATGAA 769
Db 729 GGTAAAGATTATGAATSCACACATAGGTAGTGCATTTGTTGAGAAATGTCATCAATGAG 788
Qy 770 CAC-----AAAGCCATAAGCCTTTTGAAGCGGATAACTTAATG 808
Db 789 CACAAGAAAATCTTGGCAATTTGGGAAAACATAATGAGCGTTAGGAGGTGAAGATTTAAT 848
Qy 809 GATGTTCTATTGAATCTTCAAAAAATGGAAACGTTCCAGTGCCAGTGACAAACGAAAGC 868
Db 849 GATGTTCTTAAAGACTTATGAATGATGAGGAGGCTTCAATTTCTCTATCACCAACGACAA 908
Qy 869 ATCAAGACATCCGTTTTCGAAATGTTTACTGCGGAGGAGGAAACAACTTCGAAAGCTACA 928
Db 909 ATCAAGCCATTAATTTTGGACATGTTTGTGCGGAGCAGAGACTTTCATCGTCAACAAAT 968
Qy 929 GAATGGGTAATGCGAGAGCTGATGAAATCAACTGAACCTAAGAAAAGCACAAGAA 988
Db 969 GTGTGGCTATGTTGAATGTTGAAATCCAGCGGTATTCGCGAAGCTCAACAGAA 1028
Qy 989 GTTAGACAAGTATTTGGTGAATGGGAAAGTTGATGAATCAAGATTTTATGATTTGAAA 1048
Db 1029 GTAGAGAGACATTTAGAGGAAAAGAACTTTTCGATGAAATGATGTGGAGGAGCTAAAC 1088
Qy 1049 TTTCTCAAGTTAGTGTTHAAGAAAACCTAAGATTTACATCTCCCGTT---GTCTTGATT 1105
Db 1089 TACCTAAAGTTAGTAATAAAGAAAACCTTAAGACTTTCATCCACCGTTTCCACTTTTGCT 1148
Qy 1106 CCGAGGAGGTAGAGAAACAAACGAATTTGATGATATGAATTTATCCGAAACACTCGA 1165
Db 1149 CCAAGAGATGTAGGAGAGACAAATATAACGCTACACTATTTCTGTAAAGACCAA 1208
Qy 1166 ATTGTTGTAATGCTTTGGCGATAGAGAGATCTTAATATCTTTGTCGGAACCTGGAAG 1225
Db 1209 GTCATGTTTAATGTTTGGCTTTTGGAGAGATCCAAATATTTGGAATGACGCGAAGCT 1268
Qy 1226 TTTAACCCAGAAAGTTTAAAGATTTGCAATTTGATTAATAAGGAGACGACATTTGAACTG 1285
Db 1269 TTTATGCCAGAGAGATTTGAGCAGTCTCTAAGGATTTTGTGGTAAATTTTGAATAT 1328
Qy 1286 GTACCAATTTGGTGAGAAAAGAAATATGTCCTGCGATTAATTTACAGTATTACCAATTTG 1345
Db 1329 CTTCCATTTGGTGGGAGGAGGATTTGCTCTGGGATTTGTTGGCTTAGCTAATGCT 1388
Qy 1346 GAGTATGCTAATATAAATCTATTATATCAATTTTAAATTTGGAACTGCGCGATGGAATACA 1405
Db 1389 TATTTGCCATTTGGCTCAATTAATCTATATATCACTTCGATTTGGAACCTCCCTGCTGGAATCGAA 1448
Qy 1406 CCTCAACACTTGATATGACTGAGCTATTGGCGGTCTCTCAGGAAAATAATAGATCTT 1465
Db 1449 CCAAGCGACTTGACCTTGACTAGTTGGTGGAGTAACTGCGCGCTAGAAAAAAGTACCTT 1508
Qy 1466 AAGTTGATTCCTATTCCATATCAA 1489
Db 1509 TACTTGGTTGCGACTCTTATCAA 1532
```

RESULT 15
US-11-116-881A-218
; Sequence 218, Application US/11116881A
; Publication No. US20060041949A1
; GENERAL INFORMATION:

```
; APPLICANT: Xu, Dongmei  
; APPLICANT: Nielsen, Mark T.  
; TITLE OF INVENTION: Nicotiana Nucleic Acid Molecules and Uses Thereof  
; FILE REFERENCE: 07678/141014  
; CURRENT APPLICATION NUMBER: US/11/116,881A  
; CURRENT FILING DATE: 2005-04-27  
; PRIOR APPLICATION NUMBER: 60/665,451  
; PRIOR FILING DATE: 2005-03-24  
; PRIOR APPLICATION NUMBER: 60/665,097  
; PRIOR FILING DATE: 2005-03-24  
; PRIOR APPLICATION NUMBER: 60/646,764  
; PRIOR FILING DATE: 2005-01-25  
; PRIOR APPLICATION NUMBER: 60/607,357  
; PRIOR FILING DATE: 2004-09-03  
; PRIOR APPLICATION NUMBER: 60/566,235  
; PRIOR FILING DATE: 2004-04-29  
; PRIOR APPLICATION NUMBER: 10/934,944  
; PRIOR FILING DATE: 2004-09-03  
; PRIOR APPLICATION NUMBER: 10/943,507  
; PRIOR FILING DATE: 2004-09-17  
; PRIOR APPLICATION NUMBER: 60/503,989  
; PRIOR FILING DATE: 2003-09-18  
; PRIOR APPLICATION NUMBER: 60/485,368  
; PRIOR FILING DATE: 2003-07-08  
; PRIOR APPLICATION NUMBER: 60/418,933  
; PRIOR FILING DATE: 2002-10-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 2300  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 218  
; LENGTH: 1673  
; TYPE: DNA  
; ORGANISM: Nicotiana tabacum  
US-11-116-881A-218
```

```
Query Match 23.1%; Score 400.4; DB 16; Length 1673;  
Best Local Similarity 57.7%; Pred. No. 1.2e-88;  
Matches 810; Conservative 0; Mismatches 561; Indels 33; Gaps 4;  
  
Qy 113 CCTCAGGCGCATGGAAGTTTCTTATCATAGTATATCTTCTCATTTTACTCTCTCT 172  
Db 135 CCACCAGGTCCATGGAACTACCAATCTAGAGTATGCTTCATATGTT-----TGCT 188  
  
Qy 173 GATCTAGGCCATGAACGTTTGTAGACCTTGGCTCAAAATTTATGGACCTGTTATGAGTCTT 232  
Db 189 GGACTACCAACCACTGCTCTTAGAGATTTAGCCAAAAAATATGGACCACTTATGCACCTT 248  
  
Qy 233 CAATTTGCCAAGTTTTCAGCTGTTGTCTTCTTCCAGCTGAAGCAGCCAAAGAGTTATG 292  
Db 249 CAATTAGTGAAATTTCTGCGGTTGTGTTACTTCTCTGATACGCGCAAAAGAGTATTA 308  
  
Qy 293 AAACTCAGGCTGATGCTTCGCCCAACGCCCTATCGTCTTTGGAGCGCACAGATTGTGTTT 352  
Db 309 AAACTCATGACATCGCTTTTGGCTTAGGCTAGGCCCTTTTGGCCCGAGAGATTGTCTGT 368  
  
Qy 353 TATTAATCGAAAGATGCTTGTGTTGCTTCATATGAGATCATCTGAGGCGAGATGAAGAA 412  
Db 369 TACAATAGTCTGATCTAGACCTTTTGGCCCTATGGCGACTATTGGAGACAAATGCGTAAA 428  
  
Qy 413 ATTTGGATCTTGAATTTCTGAGTCCCAAAAAAGTTTCAATCTCTCAGGTTTAAATCCGAGAG 472  
Db 429 ATATGTGTTGGAAAGTCTCAGTGCCCAAGAAATGTTCCGACATTTAGCTCTATATTAGCGG 488  
  
Qy 473 GAAGAAATGAGGATGCCATCACATTCCTCGTTCGAAAGCCGGAATCTCCGGTCAATATT 532  
Db 489 AATGAAGTTCTTCGTCTCATTAAATTTTATCCGGTCTATCTCTGTGTGAACCTATTAAATGTT 548  
  
Qy 533 ACAAGATCATTTATGGCATTTAATTTTCGATCATGATGAAGAACATCCGTTCCGTAA---T 589  
Db 549 ACGGAAAGGATCTTTTGTGTTTCAAGCTCCATGATGTAGATCAGCGTTTGGGCAAGTG 608  
  
Qy 590 TGTAAAGCAAAAGAAAGATTTGCTGAGTGTGTCGGATGCAATGAGGAGCGACGAGT 649
```

Db 609 TTCAAAGAGCAAGACAAATTTATACAACTAAATTAAGAAGTGATACTCTTAGCAGGAGG 668
QY
650 TTGGCACCAGCAGACGCTTTTCGAGCTGGAATTAATTTCACTATATCAATTTGGAGCTGAG 709
Db
669 TTGATGTGGCTGACATATTCCTTCACTGAAGTTTCTTCATGTGCTCAGTGAATGAAG 728
QY
710 TCAAAACCCAGGCGTTTGCATCAGGAGATTGACGATATATCTTTGAAGAGATTCTTTAATGAA 769
Db
729 GGTAAGATTATGAATGCACACCAATAAGGTAGATGCCATTGTTGAGAAATGTCATCAATGAG 788
QY
770 CAC-----AAAGCCATPAAGCCTTTTGAAGCGGATAACTTTAATG 808
Db
789 CACAAGAAAAATCTTGCAATTTGGAAAACTTAATGGAGCGTTAGAGGTGAAGATTTAATT 848
QY
809 GATGTTCTATTGAATCTTCAAAAAAATGGAACGTTCCAGTCCAGTGCACGACGACAAACGAAGC 868
Db
849 GATGTTCTTAAGACTTATGAATGATGAGGCGTTCAATTTCTATACCCACGACAAC 908
QY
869 ATCAAAGCATCCGTTTTCGAAATGTTTACTGCGGGAGCGAAACAACTTCGAAAGCTACA 928
Db
909 ATCAAAGCTATAATTTTTCACATGTTTCTGCTGCGGGACAGAGACTTCACTCGTCACAAAT 968
QY
929 GAATGGGTAATGGCAGAGCTGATGAATAATCCAACTGAACTAAGAAAAACA CAAGAGAA 988
Db
969 GTGTGGGCTATGTTAGAAATGGTGAATAATCCAGCCGTATTCGCGAAAGCTCAAGCAGAA 1028
QY
989 GTTAGACAGATATTTGGTGAATGGAAAGTTGATGAATCAAGATTTTCATGATTTGAA 1048
Db
1029 GTPAAGAGAAGCATTTAGAGAAAGAAACTTTTCGATGAATAATGATGTGGAGGAGCTAAAC 1088
QY
1049 TTCTTCAAGTTAGTGGTTAAAGAACTCTAAGATTACATCTCCGGTT---GTCTTGATT 1105
Db
1089 TACCTAAAGTTAGTAATAAAGAAACTCTAAGACTTCACTCCACGGTTCCACTTTTGCTC 1148
QY
1106 CCGAGGAGGTAGAGAAACAACGAATTTGATGGAATGAATAATTCACCGAACACTCGA 1165
Db
1149 CCAAGAGAATGTAGGGAAGAGACAAATATAAAGCGCTACACTATTCTGTAAAGACCAA 1208
QY
1166 ATGTTGTGAATGCTTTGGSCGATAGGAAGATCCTAATACTTTGGTCGGAACTCGAAG 1225
Db
1209 GTCATGGTTAATGTTTGGCTTTGGGAAGAGATCCAAATATTTGGAATGACCGAGAACT 1268
QY
1226 TTAAACCCAGAAAGTTTAAAGATTGTGCAATTTGATTTAATAAGGACGACATTTGAACTG 1285
Db
1269 TTTATGCCAGAGAGATTTGAGCAGTGTCTAAGGATTTTGTGGTAAATAATTTTGAATAT 1328
QY
1286 GTACCATTTGGTCAGGAAAAAGATACTCTCGCATTTACTTCAGCTATTAACCAATTTG 1345
Db
1329 CTTCCATTTGGTGGCGGAGGAGATTTGTCTGGGATTTCTGTTGGCTTAGCTAATGCT 1388
QY
1346 GAGTATGTCATTATAAATCTATTATATCATTTTAAATTTGGRACTGCCGATCGAATTACA 1405
Db
1389 TATTTGCCATTTGGCTCAATTAATCTATATCACTTCGATTTGAAACTCCCTGCTGGAATCGAA 1448
QY
1406 CCTCAAAACACTTGATATGACTGAAGCTATTGGCGGTGCTCTCAGGAAAAAATAGATCTT 1465
Db
1449 CCAAGGACTTGACTTGACTGAGTTGGTGGAGTAACCTGCCGCTAGAAAAAGTACCTT 1508
QY
1466 AAGTTGATTCCTATTCATATCAA 1489
Db
1509 TACTTGGTTGCGACTCCTTATCAA 1532

Search completed: May 31, 2006, 10:17:07
Job time : 2317 secs

This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 31, 2006, 09:38:38 ; Search time 35 Seconds
(without alignments)
5831.494 Million cell updates/sec

Title: US-10-759-813-1
Perfect score: 1733
Sequence: 1 gcataaaggaaatggagc.....tcttttcaatccgaaaaa 1733

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 246937 seqs, 5886990 residues

Total number of hits satisfying chosen parameters: 493674

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA_New.*
1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09_NEW_PUB.seq.*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US06_NEW_PUB.seq.*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_NEW_PUB.seq.*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_NEW_PUB.seq.*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10_NEW_PUB.seq.*
7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11_NEW_PUB.seq.*
8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	384.2	22.2	1751	6	US-10-953-349-22681 Sequence 22681, A
2	367.4	21.2	1504	6	US-10-953-349-24167 Sequence 24167, A
3	248.4	14.3	1737	6	US-10-953-349-14590 Sequence 14590, A
4	224.2	12.9	1915	6	US-10-953-349-13595 Sequence 13595, A
5	216.4	12.5	1817	6	US-10-953-349-2239 Sequence 2239, Ap
6	195.4	11.3	1837	6	US-10-953-349-6466 Sequence 6466, Ap
7	172.8	10.0	1958	6	US-10-953-349-17424 Sequence 17424, A
8	166	9.6	1586	6	US-10-953-349-4847 Sequence 4847, Ap
9	159	9.2	1659	6	US-10-953-349-13112 Sequence 13112, A
10	158.4	9.1	1658	6	US-10-953-349-17280 Sequence 17280, A
11	153.8	8.9	1722	6	US-10-953-349-10369 Sequence 10369, A
12	152.2	8.8	1660	6	US-10-953-349-21323 Sequence 21323, A
13	148.4	8.6	1632	6	US-10-953-349-32482 Sequence 32482, A
14	148.2	8.6	1766	6	US-10-953-349-32468 Sequence 32468, A
15	147.2	8.5	1727	6	US-10-953-349-12509 Sequence 12509, A
16	137	7.9	1639	6	US-10-953-349-32472 Sequence 32472, A
17	130.6	7.5	1711	6	US-10-953-349-33880 Sequence 33880, A
18	124.6	7.2	1720	6	US-10-953-349-32176 Sequence 32176, A
19	116	6.7	1636	6	US-10-953-349-21349 Sequence 21349, A
20	110	6.3	1852	6	US-10-953-349-12597 Sequence 12597, A
21	103.2	6.0	1667	6	US-10-953-349-2736 Sequence 2736, Ap
22	102.6	5.9	1866	6	US-10-953-349-13642 Sequence 13642, A
23	100.8	5.8	1890	6	US-10-953-349-12761 Sequence 12761, A
24	99.8	5.8	1165	6	US-10-953-349-22519 Sequence 22519, A
25	99.2	5.7	1756	6	US-10-953-349-13063 Sequence 13063, A

ALIGNMENTS

RESULT 1

US-10-953-349-22681
; Sequence 22681, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 22681
; LENGTH: 1751
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-22681

Query Match 22.2%; Score 384.2; DB 6; Length 1751;
Best Local Similarity 56.6%; Pred. No. 1.3e-92;
Matches 813; Conservative 0; Mismatches 588; Indels 36; Gaps 4;
QY 104 CAGAAATCCACCTCCAGGGCCATGGAAGTTTCTCTATCATAGGTAATCTTCTCATTTTATTA 163
DB 133 CACAAACTACCCCGGACCAAGAAGCTACCTCTGATAGGNAACCTGCATCAACTAGCA 192
QY 164 CTCACCTCTGATCTAGGCATGAAAGTTTATAGAGCTTGCGCTCAAAATTTATGACCTGTT 223
DB 193 ATGGCTGGTTTCACTTCCACATCGTACTCTCCGAGACCTTGGCCCTTAAATACGACCTCTC 252
QY 224 ATGAGCTTCAAAATGGCCAGTTTTCAGCTGTTGTTCATTTCTTACGCTCAAGCAGCCAAA 283
DB 253 ATGCATCTCCAACTTGGTGAATTTCTCCGGTGTGTATCTCTCCCAACATGGCCCAAG 312
QY 284 GAGGTATGAAAACCTCAGGCTGATCGCTTCGCCCAACGCCCTTATCGTCTTTGGACGACAG 343
DB 313 GAGATAATGAAAACCCATGATCTTCTTTTGGCAGAGCCACAAATTTCTTCTCTCTCAA 372
QY 344 ATTGTTTTTAAATCGGAAAGATCTCTTTTGTCTTATATGAGATCACTGGAGGAG 403
DB 373 ATTTTGAACCTATGGGCAAAATGATATGTTTTTGTCTTCCATATGGGATATTATGGAGCAA 432
QY 404 ATGAGAAAATTTGGATACCTTGAATTTCTGAGTGCACAAAAGTTTCAATCTCTCCAGTTA 463
DB 433 ATGAGAAAATATGTTGTCGGAGCTTCTTAAAGTGCAGAGAGTTTCACTTTTCTCTCAT 492
QY 464 ATCCGAGGAGGAGAAATGGAGGATGCCATCACATTTCTCCGTTTCGAAAGCCGATCTCCG 523

Db 493 ATTGAGAAGACGAGACATCCAAATTTATAGAATCAATTCGAATATCAAGAGTTCCACCA 552
Qy 524 GTCAATATTACAAAGATCATTTATGGCATATAATTTTCGATCATGATAAG--AACATCC 580
Db 553 ATCAATCTTACTAGTAAATTTACTCCTTGGTAAGTTCTCTGTTTCAGGGTAGCATTT 612
Qy 581 GTTGTAATTGTGAACAAAGAAAGATGCTGAGTGTGCGGATGCGAGTCAGTCAATGAGGCA 640
Db 613 GCGGACAAATCTAAGGACCAGGAGATTTTGTGTGTCTGTAAGAAATGATATTAGCG 672
Qy 641 GCGAGAGTTTGGCAGCGCAGACGCTTTTCGAGCTGCGAAATTTACTTCACTATATCATTT 700
Db 673 GGGGAGGATTTGAACCTGACGATTTGTTGTTCTTCGATGAA---ACTTCATCTCATAAAT 729
Qy 701 GGAGCTGAGTCAAAACCCAGGCGTTTGATCAGGAGATTTGAGATATATCTTGAAGAGATT 760
Db 730 GGAAGGAAGGCCAAATTTGGAGAGATGATGAGCAGGTTGACAAGATCGCAGACACATTT 789
Qy 761 CTTAATGAACACAAAGCCCAATAAGCCTT-----TTGAA 793
Db 790 CTCAGAGAGCATCAAGAAAGCGCCAAAGGCGCACTACGAGAAGGCAAGGTTGACCTAGAG 849
Qy 794 GCGGATAACTTAATCGATGTTCTATTGTAATCTTCAAAAAATGGAAGCGTTCCAGTGCCA 853
Db 850 GAGGAGGATCTTGTGTGATGTTCTTTTGAGAAATCCAAAGTGACAACTCGAGATCAAA 909
Qy 854 GTGCACAAAGCAAGCAACAAAGCATCCGTTTGGCAAAATGTTTACTGCGGAGCGCAACA 913
Db 910 ATTAGTACCAACAAATCAAGCTGTTATATTGGACGTTATTACTGCTGGAAACGATACT 969
Qy 914 ACTTCGAAGCTPACAGAATGGGTAAATGCGAGAGCTGATGAAATAATCCAACTGAACTAAGA 973
Db 970 TCAGCATCAACACTTGAATGGGCTATGGCAGAAATGATGAGAAACCCAAAGAGTGAGGAAG 1029
Qy 974 AAAGCACAAGAAAGATTAGACAAGTATTGTTGTAATGGGAAATGATGATGATCAAGA 1033
Db 1030 AAAGCACAGGCTGAAATGAACAAGCATTTAGGGAATTTGAAATAATCCATGAAACAGAT 1089
Qy 1034 TTTTCATGATTTGAAATCTTCAAGTTAGTGGTTTAAAGAACTCTAAGATTACA---TCCT 1090
Db 1090 GTAGGGAACCTTACATCTTAAGTTGGTGATCAAAAGACATTTAAGGTTTACACGCTCCT 1149
Qy 1091 CGGTTGTCTGATTCGAGGAGGTGTAGAGAAAACAACAAGTAATGATGGATATGAATTT 1150
Db 1150 TCTCCTTTATTGGTCCCTAGAGAAATGCTCTGAACCTAACCCATCATTTGATGGATATGAATA 1209
Qy 1151 CATCCGAACACTCGAATTTGTTGTAATGCTTGGCGCATAGGAGAGATCCCTAATACTGG 1210
Db 1210 CCTGTCAAGACTAAAGTCATGATAAACCGTGTGGGCAATTTGGAAGAGATCCCCCAATTTGG 1269
Qy 1211 TCGGAACCTGGAAAGTTTAAACCAGAAAGGTTTAAAGATTGTGCAATTTGATTATAAAGGG 1270
Db 1270 ACTGATGCTGAGAGGTTTGTCCGAGAGATTCGATGTTAGTCTTATCGATTTCAAGGG 1329
Qy 1271 ACGACATTTGAATCGTACCATTTGGTGAGGAAAGAAATATGCTTGGCAATTTACTTCA 1330
Db 1330 AATAACTTTGAGTATCTTCTTTTGGGCGAGGAAGAGAAATGTGCCAGGAATGACATTT 1389
Qy 1331 GCTATTACCAATTTGGAGTATGTCATATAATCTATTATATCATTTTAAATTTGGGACTG 1390
Db 1390 GGTTTAGCCCAACATTTATGCTTCCACTCGCTCTATTACTCTATCATCTTCAATTTGGGAATTT 1449
Qy 1391 GCGGATGGAATTTACACTCAACACTTTGATATGATGATGAGCTATTGGCGGTGCTCTCAGG 1450
Db 1450 CCAATGAGATGAACCTGAGGATATGGATATGTCGTGAAAACCTTTGGATTGACAGTTACA 1509
Qy 1451 AAAAAAATAGATCTTAAGTTGATTTCTTATTCATATCAAGTTAGCTTTAGGCTCAAT 1507
Db 1510 AGGAAAAGCGAGTTGTGTTTGTATTCCTCAATTTGTTAATGATCTTTGACACTGCACTGAT 1566

US-10-953-349-24167
; Sequence 24167, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; FILE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579FUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 24167
; LENGTH: 1504
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-24167

Query Match 21.2%; Score 367.4; DB 6; Length 1504;
Best Local Similarity 56.8%; Pred. No. 3.5e-88;
Matches 752; Conservative 0; Mismatches 546; Indels 27; Gaps 3;
Qy 191 TTTAGAGCTTTGGCTCAAATTTATGGACCTGTGTATGAGCTTTCAAATTTGGCCCAAGTTTCA 250
Db 9 TTAAGAGACCTTGGCATATAAATATGACCCGCTGATGATCTTCAACTTTGGACAGGTTTCA 68
Qy 251 GCTGTGTCTATTTCTTTCAGCTGAAGCAGCCAAAGAGGTTATGAAACTCAGGCTGATGCC 310
Db 69 ACTATTGTCTATTTTCATCCCTGAGTGTCTAGGGAAGTGATGAAACCCATGACATTAAC 128
Qy 311 TTCCGCCAACCCCTATGCTCTTGGACGACAGATTTGTTTATTAATAATGTTTAAATCGGAAGATGTC 370
Db 129 TTTGCCACAAGGCTTAAAGTTCTAGCTATTGAAATAATGTCTTACATTTCCACAAGTATA 188
Qy 371 TTGTTTGTCTCATATGGAGATCACTGGAGGAGAGATGAAGAAAATTTGGATCTTGAATTT 430
Db 189 GCTTTTGTCTGCTATGGAATTTATGGAGGAGCTAAGAAAAATATGCACATTTGGAGCTT 248
Qy 431 CTGAGTGCCAAAAAGTTCAATCTCTCAGGTTAATCCAGAGAGGAAGAAATGGAGGATGCC 490
Db 249 TTAAGCCTTAAACCGCTCAACTCTATTCAGCCCAATTAGAGAAGATGAGCTCTTCAATCTT 308
Qy 491 ATCACTATTCCTCCGTTCCGAAGCCGATCTCCGGTCAATATTTACAAAGATCATTTATGGC 550
Db 309 GTCAATGGATGATTCAAGAAAGGATCCCCCATCAACCTCACTGAAGCAGTACTTACA 368
Qy 551 ATTATAATTTTCGATCATGATAAGAACATCCGTTGG---TAAATTGTAAGCAAAAAGAGA 607
Db 369 TCAATTTTACAATAATTTGCTTCAAGGGCTGCCCTTTGGCAAGAAATTGCAAGACCAGAAAA 428
Qy 608 TTGCTGAGTGTTCGCGATGCGAGTCAATGAGGAGCGACGAGTTTGGCACCCGACGCGCT 667
Db 429 TTTATATCAGTGTGTTAAAAAACAATCAAAAACCTTGTGCAAGGTTTGGCAATTGAAGATTG 488
Qy 668 TTTCCGAGCTGGAATTTACTTCACTATATCATTTGGAGCTGAGTCAAAAACCCAGGCGTTTG 727
Db 489 TTTCTCTTCTGTTACTTGGCTTCAACATGTCTACTGCTTGGAGGCTAAGCTTGAAGGTTG 548
Qy 728 CATCAGGAGATTGACGATATACCTTGAAGAGATTCTTAATGAACAACAAGCCCAATAGCCT 787
Db 549 CATCAACAGGCTGATCAGATAATGAAAAACATCATCAATGAGCATAAAGAGGCAAAATCG 608
Qy 788 TTTGAAGCGGATAAATCTTAA-----TGGATGTTCTTATGATCTT 826
Db 609 AAAGCAAGGATGACAAAAGTGAAGCAAGAAGAGATCTTGTGGATGTTCTCATACAATAT 668
Qy 827 CAAAAAATGGAACGTTTCCAGTGCCAGTGACAAAACGAAAGCATCAAGCATCCGTTTGG 886
Db 669 GAGGATGGAGAGCAAGGATTTTCTTAACTAGAAACAGATCAAGGCCATTAATCTCG 728
Qy 887 CAAATGTTTACTCCCGGAGCGAAAACAACTCCGAAGCTTACAGAAATGGGTAAATGGCAGAG 946
Db 729 GACATTTTGTCTGTGGAGTGAGACACAGCAACCAACCATAGATTTGGGCAATGGCTGAA 788

QY 947 CTGATGAAATCCAACTGAACCTAAGAAAGCAAGAAAGCTTAGACAAGTATTGGT 1006
 DB 789 ATGGTAAAGATCCACAGTAAAGAAAGCAACATCTGAGTGAGAGGTATGCAAT 848
 QY 1007 GAAATGGGAAAGTGTAGTAAATCAAGATTCATGATTTGAAATCTTCAAGTTAGTGGTT 1066
 DB 849 ATGAAAGCAAGGTTGATGAAATTCGATCAATGAACTTCAATATTTGAAACTTAAATG 908
 QY 1067 AAGAAACTCTAGATTAACATCTCCGG---TTGCTTGTATTCGGAGGAGTGTAGAA 1123
 DB 909 AAGAGACCTTGAGGTGACCCCTCCAGCTCTCTTTTGTGCTCAAGAGAATGTGTCAA 968
 QY 1124 ACAACACGAATGATGGATATGAAATTCATCCGAACACTCGAATTTGTGTGAATGCTTGG 1183
 DB 969 ACATGTGAGATACATGTTATCACATACCAGCCAAACCAAGTCAATGCTGCTGG 1028
 QY 1184 GCGATAGGAGAGATCCTTAATCTTGGTGGAACTCTGGAAAGTTTAAACCAAGAGGTTT 1243
 DB 1029 GCAATTTGGAAGAGATCCAACTATTTGAGTGAATCAGAGAGGTTTATCTGAGAGGTTT 1088
 QY 1244 AAGATTGTGCAATGATTAAGAGGAGCAGCAATTTGAATGTTGATGTTGTTGTCAGGA 1303
 DB 1089 ATTGATGACATATTGACTACAAAGGAGTAATTTTGGTTCATCTCTTTTGGTGTGCTGA 1148
 QY 1304 AAAAGAAATATGCTCGCATTTACTTCAGCTATTACCAATTTTGGAGTATGTCATTATAAAT 1363
 DB 1149 AGAAGATATGCGCAGAGACATTTGCTTTGAGAGCTCGGAACCTAGCCCTTGCATG 1208
 QY 1364 CTTATATATCATTTTAAATTTGGGAACTGGCGGATGGAATTTACCTCAACACATTTGATG 1423
 DB 1209 TTGTTGTATCACTTTGATGGAAGCTTCCAAAGTGAATGAGAAGTGGAGAACTGGACATG 1268
 QY 1424 ACTGAAGCTATTGGCGTCTCTCAGGAAAGAAATAGATCTTAAAGTTGATCTCTATTCCA 1483
 DB 1269 AGTGAAGATTTGGAGTCAACTATTAAGAAAGATAATCTATTCTTTGTTCTCTTTCCA 1328
 QY 1484 TATCA 1488
 DB 1329 TATCA 1333

RESULT 3
 US-10-953-349-14590
 ; Sequence 14590, Application US/10953349
 ; Publication No. US20060107345A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ALEXANDROV, Nikolai et al.
 ; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
 ; FILE OF INVENTION: ENCODED THERBY
 ; FILE REFERENCE: 2750-1579PUS2
 ; CURRENT APPLICATION NUMBER: US/10/953,349
 ; CURRENT FILING DATE: 2004-09-30
 ; NUMBER OF SEQ ID NOS: 40252
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 14590
 ; LENGTH: 1737
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; US-10-953-349-14590

Query Match 14.3%; Score 248.4; DB 6; Length 1737;
 Best Local Similarity 51.4%; Pred. No. 1.1e-56;
 Matches 722; Conservative 0; Mismatches 646; Indels 36; Gaps 5;

QY 48 TAAATAGTTTCTGCTGTTTAAATCTTAGTAGTAGTCAATGAGGTTGTGGAAGAAACAGA 107
 DB 171 TAACTCTTCTTTGTTGTTCTCTTCAATACCGTAAAAACCTTCAAGAAATCCAC 230
 QY 108 ATCCACCTCCAGGCCATGGAAGTTTCCCTATCATAGGTAATCTTCTCTATTATTACTCA 167
 DB 231 CACTTCCACGAGTCTAGAGGCTTCCCAATAATAGGGAATCTTTCATCAACTAGATAGTC 290

QY 168 CTTCTGATCTAGGCCATGAACGTTTTAGAGCCTTGGCTCAAAATTTATGACCTGTTTATGA 227
 DB 291 CTTCTCTT-----CATGAGCAGCTATGCGATCTCTCAAAGAAATATGCTCTTTATTTT 344
 QY 228 GTCTTCAAAATGGCCAAAGTTTTCAGCTGTTTGTTCATTTCTTTCAGCTGAAGCAGCCAAAGAGG 287
 DB 345 CCCTTCAATTTGGATTAAGGCCAGCCATAGTTGTTTCTCTCCAAATTTGGCCAAAGAGG 404
 QY 288 TTATGAARAACTCAGGCTGATGCTTCCGCCCAAGCCCTATCGCTTTGGAGCGCACAGATTG 347
 DB 405 TAATGAAGACAATGACCTTGAGTGTGTGGAGCAGCTAAATTTACTAGGCCAACAGAAAC 464
 QY 348 TGTTTTATAATCGGAAAGATGTCTTGTGTTTCTTATATGGAGATCACTGGAGCGACAGATGA 407
 DB 465 TAACTTATAATGGTTAGACATGGGATTTTCTCATACGATAGTTATTTGGAGAGAATCA 524
 QY 408 AGAAATTTTGGATATCTTGAATTTCTGAGTGCCAAAGAGTTTCAATCTCAGGTTTAAATCC 467
 DB 525 GAAATTTTGTGTTGTCATGCTTAGCTCCAAAGCTGTCCAAAGCTTTACCTCAATAA 584
 QY 468 GAGAGGAAGAAATGGAGGATGCCATCACATCTCCTCGTTTCGAAAGCGGATCTCC----- 522
 DB 585 GACATTTTGAAGTCAAGCAGATGATAAAAAAATATCTAGGCATGCTCATCTTCAAAAG 644
 QY 523 -GGTCAATATTACAAAGATCATTTATGGCATTTATAATTTTCGATCATGATAAGAACATCCG 581
 DB 645 TTACAAATTTAAATGAAGTCTTATATCCCTTATTAGTACTATTGTATGTAGAAATTTGTTT 704
 QY 582 TTGGTAAATTTGAAGCAAAAGA-----AAGATTGCTGAGTGTTCGCGATG 626
 DB 705 TGGGAGAGAGGTATGAAGAGGAGGATCTGAAGAGTAGCAGATTCACATAAAGTTGTTCAATG 764
 QY 627 CAGTCAATGAGCAGCGAGGTTTGGCACCGCAGACGCTTTCCGACGCTGCGAAATTAAC 686
 DB 765 AGTGTGAAGCTATGTTGGTAACTCTTCTTTGTTTTCAGATATTATTCCTTTTCATGGGTTGGA 824
 QY 687 TTCCTATATCATTTGAGCTGAGTCAAAACCCAGGCGTTTTCATCAGAGGATTTGACGATA 746
 DB 825 TTGATAAACTCAGGGGACTGGATGACGCTTTGAAAGCAATTTCAAGAGAGATGGAATAGT 884
 QY 747 TACTTGAAGAGATTTCTTAATGAACACA-----AAGCCAAATGAAGCCTTTTGAAGCGGATA 800
 DB 885 TCTACCAAGAGGCCATTTGATGAACACATGAATTTCTAAAGAAAAAATCTCCAGAGGAAAGAG 944
 QY 801 ACTTAAATGAGTGTCTTATTGATCTTCAAAAAATGGAACCTTCCAGTCCGAGTACACA 860
 DB 945 ATTTAGTTGATGCTTTACTTCAACTGAAAGAGAACAAACGTTTCCCATAGATCTCACCA 1004
 QY 861 ACGAAAGCATCAAAAGCATCCGTTTTTGCATAATGTTTACTCCGGGAGCGAAACAACTTCGA 920
 DB 1005 ATGATAACATCAAGCAGCTGCTTTGAACTTACTTTGAGGACTACAGGTACAACTGAAG 1064
 QY 921 AAGCTACAGAAATGGGTAAATGGCAGAGCTGATGAAAAATCCAACTGAACCTAAGAAAAAGCAC 980
 DB 1065 TCACAAACAATCTGGGCCATGACTGAACATAAATAAAAAATCCAAAGCATTTATGAAGAAAAAGTGC 1124
 QY 981 AAGAAGAAAGTTAGACAAGTATTGTTGTAATTCGGAAAAAGTTTTCATGAATCAAGATTTTCATG 1040
 DB 1125 AAGAAGAAATTAGAGGCTTTAAGTGTAAAGAAAGATTTCTTGGATGAAGATGATATTACAA 1184
 QY 1041 ATTTGAAATTTCTTCAAGTTTAGTGGTTTAAAGAAACTCTAAGATTACA---TCCTCCGGTTG 1097
 DB 1185 AGTTTTCCTACTTAAAGGCGAGTGAATAAAGAGACACTTAGATTGCACTTTACCAGCACCAC 1244
 QY 1098 TCTTGAATTCGAGGAGGTGTAGAGAAACAAACAGAAATTTGATGATATGAAATTTTCATCCA 1157
 DB 1245 TACTTATACCAAGAGAGAAACAAATAAAAAATGCAATGTTAGTAGGCTTACGAAATTTCCAGCCA 1304
 QY 1158 ACACCTGAATTTCTTGAATGCTTGGCGATAGGAGAGATCCTTAATATCTTCTGTCGGAAC 1217
 DB 1305 AGACATTTATATATGTAATGCTTGGGCAATTTCTATAGAGATCTTAAAGGCTTGGAAAGATC 1364
 QY 1218 CTGAAAGTTTAAACCCGAAAGGTTTAAAGATTGTGCAATTTGCAATTTATAAAGGAGCAGCAT 1277

Db 1365 CAGAGAGTTTATACCTGAGAGGTTTTAAATTTGGACATAGATCTTTATGGACAAGATT 1424
Qy 1278 TTGAACCTGGTACCATTTGGTGCAGAAAAGAAATATGCTCGCATTTACTTCCAGCTATTA 1337
Db 1425 TTGAGTTTATTCATTCGGTGTGCTGAGATTGCGCCTGGTATGATATGATGGCATTTG 1484
Qy 1338 CCAATTTGGAGTATGCTATTAATAATCTATTAATCAATTTTAAATTTGGAACTGGCCGATG 1397
Db 1485 CTGCATTGGATCTTATCTATCTGTAATCTCTTTATCTTTGACTGGGAATTTGCCCAAG 1544
Qy 1398 GAATTACACCTCAACACCTTGATA 1421
Db 1545 GAATGAAAAGGAAGACATTGATA 1568

RESULT 4

US-10-953-349-13595
; Sequence 13595, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 13595
; LENGTH: 1915
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-13595

Query Match 12.9%; Score 224.2; DB 6; Length 1915;
Best Local Similarity 51.2%; Pred. No. 38-50;
Matches 762; Conservative 0; Mismatches 668; Indels 57; Gaps 8;
Qy 48 TAATAAGTTTCTCGCTGTTGTTTAAATCTTAGTAGTAGTCATGAGGTTGTGGAGAAACAGA 107
Db 65 TAAACAGTTTCTCTGCATCTTCTTATATCTTTCAGCTAATAGAAGAAACAAATACA 124
Qy 108 ATCCACCTCCAGGGCCATGGAAGTTTCTATCATAGGTAATCTTCTCATTTATTAATCA 167
Db 125 ATCTGCCACCATCCCAACCAAGATACCCATAATCGCAATCTTCCACGCTA----- 177
Qy 168 CTTCGTATCTAGCCATGAACGTTTATAGAGCCTTGGCTCAAAATTTATGACCTGTTATGA 227
Db 178 --GGCACACTGCCACACCCGCTCTTTCATGCACTCTCACAAATATGCGCCCTCTCATGA 235
Qy 228 GTCTTCAAAATGGCCAAAGTTTCAGCTGTTGTCAATTTCTTTCAGCTGAAGCAGCAAGAGG 287
Db 236 TGTTCGAATTTGGTCAAAATTCCAACCTAGTGTCTCATCAGCTGACGTTGGCCAGAGAA 295
Qy 288 TTATGAAAACTCAGGCTGATGCTTCGCCCAACGCCCTATCGTCTTGACGCGACAGATG 347
Db 296 TAATCAAAACGCGATGATGTTGTTTCTTCCAAACCGCCGCAACCTACAGCTGCTAAATCT 355
Qy 348 TGTTTTATAATCGGAAAGATGCTTTGCTTTCATATATGAGATCACTGGAGGCGAGATGA 407
Db 356 TTGGTTATGAGTACAAAGATGCGGCTTTCGTGTACTACCGCAAGAGTGGAGACAAAGA 415
Qy 408 AGAAAAATTTGGATCTTGAATTTCTGAGTGCCAAAAAGTTTCAATCTCCAGGTTAATCC 467
Db 416 TAAAGACATGTAAGTTGAGCTTATGAGTCTGAAGAGTGCGGTTGTTTCATTCAATTA 475
Qy 468 GAGAGA-----AGAAATGGAGGATGCCATCACATCTCTCCGTTGGAAAGCCGGAT 518
Db 476 GACAAAGAGTTGTTACAGAGTTGGTTGAAGCTATAGGTGAACGCTGTGGTAGTGAAGAC 535
Qy 519 CTCGCGTCAATATTCAAGATCAT-----TTATGGCATTAAT 558

RESULT 5

US-10-953-349-2239
; Sequence 2239, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.

Db 536 CATGTGTGAATCTGACTGAGATGCTGATGGCAGCATCGAACGACATTTGTGTCTAGATGTG 595
Qy 559 TTGATCATGATAAGAACATCCGTTGGTAAATTTGTAAGCAAAAAGAAAGATTTGCTCAGTGT 618
Db 596 TTCTTGGACGGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 655
Qy 619 TGCCGATGTCAGTCAATGAGGC-AGGACGAGTGTGGCAGCAGCAGCAGCAGCAGCAGCAG 677
Db 656 TGGGAAGAAAGATTTATGAGACTATATACGGCTTTTACGGCTTTGAGGCTGGGTGATTTCTCCCTTGT 715
Qy 678 GGAATTAATCTCACTATATCATTTGAGCTGAGTCAAAACCCAGGCGTTTGCATCAGGAGA 737
Db 716 TGGGTTGGTTGACTATCTGACTGGCTTAATTCAGAGATGAAACACGCTTTCTCGCAG 775
Qy 738 TTGACGATATATCTTCAAGAGATTTCTTAATGAACACAAAGCCAAATGAAGCCTTTTGAAGCG 797
Db 776 TAGATGCTTTCTTGTGATGAGGTAATTCAGAACACGAGAGCAGTAAACAAGAA--TG 832
Qy 798 ATAACTTAATGATGTTCTTATTTGAATCTTCAAAAAAATGGAACGTTTCCAGTGCCAGTGA 857
Db 833 ATGACTTCTTGGGATACTTCTTCAACTTCAAGAAATGTGGAGGCTTGACTTTTCAGCTCG 892
Qy 858 CAAACGAAGCATCAAAAGCATCCGTTTGTGCAAAATGTTTACTGCGGAGCGAGCAAACTT 917
Db 893 ACCGAGATTAACCTCAAGCAATCTTAGTGGACATGATATAGTGGGAGTGACACTACTT 952
Qy 918 CGAAAGCTACAGAATGGGTAAATGGCAGAGCTGATGAAAAAATCCAACTGAACTAAGAAAAG 977
Db 953 CAACAACTCTAGAATGGACTTTTGGGAGTTCTTTAGAAATCCAAATACCATGAAGAAAAG 1012
Qy 978 CACAAGAAGAAGTTAGACAAGTATTGG-----TGAATGGAAAAAGTTGATGAATCAA 1031
Db 1013 CTCAGAAGAGGTAAAGAGGTGGTGGGAATCAATTTCCAAAGCAGTACTGGATGAAAAAT 1072
Qy 1032 GATTTCATGATTTGAAATTTCTTCAAGTTAGTGGTTTAAAGAACTCTAAGATTTACATCTCT 1091
Db 1073 GTGTGAATCAATGAACACTTGTGAAATGTGTAGTCNAAGAAACTTTGAGATTTACATCCAC 1132
Qy 1092 CGGT---TGTCTTGAATCCGAGGGAGGTAGAGAAACCAACGAAATTTGATGGATATGAAA 1148
Db 1133 CCTTCTCTTTTGTGCTCGAGAGACATCATCAAGTGTAAACTTAAGAGGGTACGATA 1192
Qy 1149 TTATCCGACACACTCGAATTTGTTGATGCTTGGCGCATAGGAGAGATCCTAATACTT 1208
Db 1193 TTCCCGCAAAACCAATGGTATTTATCAATGATGCGCGATCCAGAGGATCTTGAATAT 1252
Qy 1209 GGTGGACCTTGGAAAGTTTAAACCCAGAAAGGTTTAAAGATTTGTGCAATTTGATTAAG 1268
Db 1253 GGGATGATCTTGAAGATTTATTTCCGAAAGATTTGAAACTAGCCAACTGATCTTAATG 1312
Qy 1269 GGACGACATTTGAACTGGTACCAATTTGGTGCAGGAAAGAAATATGCTCTGGCATACTT 1328
Db 1313 GACAAGATTTTCAATTAATTCGGTTCCGTTATGGGAGAAAGGGATGCGCTGCAATGTCTAT 1372
Qy 1329 CAGCTATTACCAATTTGGAGTATGCTATTAATAATCTATTAATCATTTTAAATTCGGAA 1388
Db 1373 TTGACTTGGCTTCAACTGAGTATGTTTGTGTAATCTTTTGTATTTGGTTCAATTTGGAATA 1432
Qy 1389 TGGCGCA-----TGGAAATTACACCTCAAAACACTTGTATGATGACTGAAGCTATTGGCGGTG 1442
Db 1433 TGTCCGAGTCTGGAGCTATATTTGATGCACAACTTGACATGATGAGTGAAGCAATGACTCA 1492
Qy 1443 CTCACAGAAAAAATAGATCTTAAGTTGATTTCTTATTTCCATATCAA 1489
Db 1493 CTGTGAGTAAAGAAAGTACCACCTTCATCTTTGAACCAAGAACCATATAAA 1539

Db 147 TACATCAGATTGGAGAAATGGCTCA CAGGTCACATTTCAACAT-----CTCGCGGAAA 197
Qy 210 TTTATGGACCTGTTATGAGTCTTCAAAATTTGGCCAAGTTTTCAGCTGTGTCATTTCTTCAG 269
Db 198 GATACGGACCTGTGATGCTTCTTCACTTTGGTTTGTCCCTATAACTGTGGTCTCATCGA 257
Qy 270 CTGAAGCAGCCAAAGAGGTTATGA AAAA CTCAAGGCTGATGCCCTTCCGCCAAAGCCCTATCG 329
Db 258 GAGAGCCGCTCAAGAAGTGTAGAACTCATGACCTAGACTGTGTG CAGCAGGCTTAAGC 317
Qy 330 TCTTGGACGCACAGATGTGTTTTATAATCGGAAGAGATGTCTTGTTTGCCTTCATATGAG 389
Db 318 TTGTCGGACAAAGGTTACTCTCGCGGATTTTAAAGATATCGGTTTTTACGCCATACGGTA 377
Qy 390 ATCACTCGAGGCAGATGAAGAAAATTTGGATACTTTGAATTTCTGAGTGC CAAAAGTTTC 449
Db 378 ACGAGTGAAGCGCGGGGTAAAGTTTGGCTCGCTGAGCTTTTCTGTTTGA AAAAGGTTTC 437
Qy 450 AATCTCTCCAGGTTAATCCGAGAGGAAGAAATGGAGGATGCCATCACATTCCT-----CC 503
Db 438 AGTCCTTTAGGCATATCCGAGAGGAAGAAATGTAATCTTCTGCTCAAGCAACTGTCGGAAT 497
Qy 504 GTTCGAAAGCCGATCTCGGTCAATATTAACAAGATCATTTATGCGCATTTAATTTTCA 563
Db 498 CTGCGGTTGATCGCTCTCGGTGATTTGAGCAAAATCCCTTTTCTGGCTAACCGCTAGTA 557
Qy 564 TCATCATGAACATCCGTTG-----TGAGTGGT-----TAATTTAAGC 596
Db 558 TCCTTTTTAGAGTTGCTTTAGGACAGAAATTTTCAAGAGCGATTTTATCGATAAGAAA 617
Qy 597 AAAAAAGAAAGATTGTGAGTGTGCGGATGCGATCAATGAGCAGCAGGATTTTGGCA 656
Db 618 AGATCGAAGAGCTCGTGTGCGAAGCTGAGACTGCCCTAGCAAGTTTCACTGTTCTGATT 677
Qy 657 CGCAGACGCTTTCCGAGCTGGAATTAATTCATATATCATTTGAGCTGAGTCAAAAC 716
Db 678 TCTTCCCTGTTGCGGAGCTGAGTGGCTCGTTGATTGGTTTTCCGGAACAACAAGAGAC 737
Qy 717 CCAGGCTTTGTCATCAGGAGATTGACGATATATTTGAAGAGATTCTTAATGAACAACAAG 776
Db 738 TCACAGATGTTTTTTACAGCTCGATGCTCTGTTTCAACATGTCATAGATGATATTAA 797
Qy 777 CCAATAAGCCTTTTGAAGCGGATAAATTAATGATGTTTCTATTG-----AATCTTC 827
Db 798 ATCTCGAAGATCAAAAGAGCAGCAAGACATCATCGATTCAATGTTGGATGTGATTCTATA 857
Qy 828 AAAAAATGGAACGTTTCCAGTGCAGTGACAAACGAAGCATCAAGCATCCGTTTTC 887
Db 858 AACAAAGGAGAGACAGTTCTTTAGAGCTCACAAATAGATCATATCAAGGGGTTTCTCGGCA 917
Qy 888 AATGTTTACTCGCGGAGCGGAAAACAACTTCGAAAGCTACAGAAATGGGTAATGGCAGAGC 947
Db 918 ATATATTTCTGACGGGATAGACACAGGGGCCATCACCATGATATGGGCAGTGAACGAGC 977
Qy 948 TGATGAAAAATCCAATGAACTAAGAAAAAGCAACGAAGAAGTTAGACAAGTATTG----- 1004
Db 978 TCGTTAAAAAACCCGAACTGATAAAGAAAGTTTCAAGCGGATATCCGAGAACAACTTGGCA 1037
Qy 1005 GTGAAATGGGAAAGTTGATGAATCAAGATTTTCATGATTTGAAATTTCTCAAGTTAGTGG 1064
Db 1038 GCAATAAGGAGAGAAATCACCGAGGAAGATATCCAGAAAAGTTCTTACTTTTGAAGATGGTAA 1097
Qy 1065 TTTAAGAAATCTTAGATTACATCTCTCGGTTG---TCTTGATTCCGAGGAGTGTAGAG 1121
Db 1098 TCAGAGAAACATTCAGGTTTACCCAGCAGCTCCTCTTATATCTTCCAGGGGAAACAATGG 1157
Qy 1122 AAACAACACGAATTGATGATATGAATTCATCCGAACACTCGAAATTTGTTGTAATGCTTT 1181
Db 1158 CTCACATCAAAAGTTTCAAGGGTATGATATTCTCTCCCAAGAGGAGGATCTTGGTCAATGTTT 1217
Qy 1182 GGGCGATAGGAGAGATCTTAATACTTGTGTCGGAACTCGAAAGTTTAAACCCAGAAAGGT 1241

Db 1218 CGCAATAGGAAGAGATCCCAAACTCTGGACAAACCCGAAAGAGTTTTCACCTCGAGAGGT 1277
Qy 1242 TTAAAGATTTGCAATTTGATTTATAAAGGGACGACATTTTGAACTGGTACCACTTTTGGTGAG 1301
Db 1278 TTATGGATAGCTTTGTTGATTTATAGGGACAAACATTACGAGCTCTTACCACTTTGGGTCCG 1337
Qy 1302 GAAAAAGAAATGCTCTGGCAATTACTTACGCTATTAACAATTTGGAGTATGTCATTATAA 1361
Db 1338 GTCGAAGGATATGTCCCGGATGCGCAATGGGGATTTGTCGCGTCGAAATTTGGACTCTTGA 1397
Qy 1362 ATCTATTATATCATTTTAAATTTGGGAACCTGGCGGATGGAATTTACACCTCAAAACACTTGATA 1421
Db 1398 ACTTACTTTTCTTTCGATTGGAAGTTTGCCTGATGGATGCACATATAAAGATATCGATA 1457
Qy 1422 TGACTGAAGCTATTGGCGGTGCTCTCAGGAAAAAATAAGATCTTAAAGTTGATTCTCTATTC 1481
Db 1458 CTGAAGAAGCTGGTACTCTTACAATAGTCAAGAAAGTACCTCTCAAGCTCGTTCAGTTTC 1517
Qy 1482 CATATCA 1488
Db 1518 GAGTTCA 1524

RESULT 7
US-10-953-349-17424
; Sequence 17424, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953.349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 17424
; LENGTH: 1958
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-17424

Query Match 10.0%; Score 172.8; DB 6; Length 1958;
Best Local Similarity 57.3%; Pred. No. 1.2e-36;
Matches 353; Conservative 0; Mismatches 257; Indels 6; Gaps 2;

Qy 877 ATCCGTTTTCGAAAATGTTTACTGCGGGAGCGGAAAACAATTCGAAAGCTACAGAAATGGGT 936
Db 1185 ATATTTTCAGGATATGTTTGTAGCAGGAAGTGAACAACCGCATCAGCACTAGAAATGGGC 1244
Qy 937 AATGCGAGCTGATGAAAAATCCAACTGAACCTAAGAAAAGCACAAGAAAGATTAGACA 996
Db 1245 TATGCGAGCTCATGAAAAATCCAAATGAAATTTGAAGAAAGCCAGGATGAGGTGAGAAA 1304
Qy 997 AGTATTTTGGTGAATGGGAAAAGTTGATGAATCAAGATTTTCATGATTTTGAAATTTCTTCAA 1056
Db 1305 ATTTGTAGGGAATAATCAAAAGTAGAGAAAATGACATCAATCAATGACATACATGAA 1364
Qy 1057 GTTAGTGTGTTAAAGAAACTCTAAGATTACATCCTCCGG---TTGCTTTGATTTCCGAGGGA 1113
Db 1365 ATGTGTAATCAAGAAACTCTAAGGTTACATCCACCTGCTCCTCTCTTGGCTCCTTAGAGA 1424
Qy 1114 GTGTAGAAAACAACACGAAATTTGATGGAATTAAGAAATTCATCCGAACACTCGAATTTGTTGT 1173
Db 1425 GACAGCATCTAGTGTGAAAACCTGGGAGGGTATGATATTCAGCAAAAAACA CTGGTATATGT 1484
Qy 1174 GAATGCTTTGGCGATAGGAGAGATCCCTAAATACTTTGGTCGGAACCTGGAAAGTTTAAACCC 1233
Db 1485 CAATCGATGGGCAATTCAGAGGACCCTGAGGTTTTTGGGAAAGACCTGAAGAGTTCAATCC 1544
Qy 1234 AGAAAGGTTTAAAGATTGTGCAATTTGATTATAAAGGGACGACATTTGAACCTGGTACCATTT 1293
Db 1545 AGAGAGACATGATAACAGCCGGGTTTCATTTTAAATGCCCAAGACTTGCATTTTATTACATTT 1604

QY 1294 TGGTGCAGGAAAAAGAAATATGCTCGCATTACTTCAGCTATTACCAATTTGGAGTATGT 1353
 DB 1605 TGGTTTGGAGAGGGCATGCCCGAATGACGTTTGGCTTGTCTGTTGAGTATAT 1664
 QY 1354 CATTATAAATCTATTATATCAATTTTAAATGGGAACGCGCATGGAATTAACACCT---CA 1410
 DB 1665 TCTGGCCAACTTCTTTATTGGTTCAATTGGAAGCTGCCTGCAACTCACACATCTGGACA 1724
 QY 1411 AACACTTGATATACATGAGCTATTGGCGTGCTCTCAGGAAAAAATAGATCTTAAGTT 1470
 DB 1725 AGACATAGACATGAGTGAACATATGGCTTAGTTACCTACAGAAAGAGCACTTCATCT 1784
 QY 1471 GATTCCTATTCCATAT 1486
 DB 1785 TAAACCAATCCCCTTT 1800

RESULT 8

US-10-953-349-4847
 ; Sequence 4847, Application US/10953349
 ; Publication No. US20060107345A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ALEXANDROV, Nikolai et al.
 ; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
 ; FILE REFERENCE: 2750-1579PUS2
 ; CURRENT APPLICATION NUMBER: US/10/953,349
 ; CURRENT FILING DATE: 2004-09-30
 ; NUMBER OF SEQ ID NOS: 40252
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 4847
 ; LENGTH: 1586
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; US-10-953-349-4847

Query Match 9.6%; Score 166; DB 6; Length 1586;
 Best Local Similarity 49.1%; Pred. No. 6.9e-35;
 Matches 705; Conservative 0; Mismatches 680; Indels 51; Gaps 8;

QY 2 CATAAAGAGAAATGGAGCAGAAAAATCTCTCTTTCCGAGCATTTTAAATAGTTTCTG 61
 DB 17 CAAAAACAAAAACAGAGCGAGGAATAATGTCACTCTCTCTCTCTCTCTCTCTCTCT 76
 QY 62 CTGTGTTTAACTCTTAGTAGTATCATGAGTTGTGGAAGAAACAGAAATCACTCCAGG 121
 DB 77 CCCCTTATCTTAATCTCTTGAANAATCTCAAAACCATCGAATGGAAGCTTCTCCAGG 136
 QY 122 CCATGGAAGTTTCCATCATAGTAAATCTCTCTCTATTTATCTCACTTCTGATCTAGG 181
 DB 137 CCAAGAGAGTTCCGATCATCGGAA-----CTTACACCAACGCCGGGAATTATCTC 189
 QY 182 CATGAACGTTTATAGCCTTGCTCAAAATTAATGACCTGTTATGAGTCTTCAAAATGGC 241
 DB 190 CAGGAACAGTCGGAATCTTT--CCGAAAAGTAGCGACCAATCGTGTCTTCCGATACGA 247
 QY 242 CAACTTACGCTGTTGTCAATTTCTTCAGCTGAAGCAGCAAGAGAGTTATGAAAATCAG 301
 DB 248 TTCGTCCTCGTGGTGTGATCTCGTCAAAAGAGCAGAGGAAGTTCTCAAGACCCAC 307
 QY 302 GCTGATGCTTCGCCCAACGCCCTATCGTCTTGGACGACAGATGTGTTTTATATCGG 361
 DB 308 GATCTTGATGTTGTAGCCGACAGAGAGCGTTGGACACGAGCAATCTCTTACACTTT 367
 QY 362 AAAGATGCTTGTGTTCTTATATGAGATCATGAGGACGATGAGAAAAATTTGGATA 421
 DB 368 AAAGACATCGGATTCGACCGTACGTTGAAGATTGGAGAACGATGCGGAAGCTCTCGGTG 427
 QY 422 CTTGAATTTCTGAGTCCAAAAGTTCATCTCCAGGTTTAAATCCGAGAGCAAGAAATG 481
 DB 428 GTCGAGCTCTTCAGCTCGAAAAAGCTTCAATCTTTCAGGTATATCAGAGAGGAAGAAC 487

QY 482 GAGGATGCCATCATCTTCTCGTTCGAAAGCG-----GATCTCGGTCAATATTACA 535
 DB 488 GACTTGTGTCAAGAACTCTCTGATTTAGCTTCGAGAGCATCTTTGGTGAATCTTGAG 547
 QY 536 AAGATCATTTATGGCATTATAATTTTCGATCATGATAAGAACATCCCGTTGCTTAATTTGAAG 595
 DB 548 AAAACCCCTTTTACTTTTAGTCGGAAGTATAGTGTAGGATAGGGTTTGGATAAATCTC 607
 QY 596 C-----AAAAAGAAAGATTCTCGAGTGTTCGCCGATGCGAGTCAGTCAATGAG 637
 DB 608 CGTAGTGTGAGTTCTGATGAAGATAGCATCGATGATCTCTGTGCACAGTCTGAAGAT 667
 QY 638 CGAGCCAGGATTTTGGCACCGCAGACGCTTTTCC-----GAGGTGGAATTTACTTTCAC 691
 DB 668 GTCATAAGGAATCTATCTCTGATTTCTTCCCGGATTAATGGGTAGGCTCATCGAG 727
 QY 692 TATATCATTGGAGCTGAGTCAAAACCCAGGCGTTTGCATCAGGAGATTGACGATATATCT 751
 DB 728 TGGATCTTCAGTGAGAGGAAGAGATTGAATAGACTTTTACTCGGAAGTAGACACTTTCTTT 787
 QY 752 GAAGAGATTTCTTAATGAAC-----ACAAAGCCAATAAGCTTTTGAAGCGGATACTTA 805
 DB 788 CAGAATATTTCTTGATGATCATCTCAAGCCTGGAAGAGAGAGCTCTGATATCATCGATGT 847
 QY 806 ATGGATGTTCTATTGATCTTCAAAAAAATGGAACGTTCCAGTGCAGTGACAAACGAA 865
 DB 848 ATGATTGATGATGAGAGCAAGAGAAAGAGGAGACTCTTTCAAGTTCCACCTGAT 907
 QY 866 AGCATCAAAAGCATCCGTTTTCGAAATGTTTACTGCGGAGCGAACAACACTTCGAAAGCT 925
 DB 908 CATCTCAAAAGGATGATCTCGACATATTTCTAGCAGGAGTTGGAACAAGCTCCACTACA 967
 QY 926 ACAGATGGGTAAATGGCAGAGCTGATGAAAAATCCAACTGAACATAAGAAAGACAAAGAA 985
 DB 968 TTGATATGGGCGATGACCGAGCTGATCAGAAAACCCGAGAGTGTATGAAGAAGTTCAAGAC 1027
 QY 986 GAAGTTTAGACAAGTATTTGGTCAAAAT---GGGAAAGTTGATGAATCAAGATTTTCATGAT 1042
 DB 1028 GAGATTCGACAACTTTGGGACAGAGAGAGAGATCACAAGAGAAGATTTAAACCAG 1087
 QY 1043 TTGAAATTTTCAAGTTAGTGTGTTAAAGAACTCTAAGATTACATCTCCCGTT---GTC 1099
 DB 1088 CTCTACTACTTCAAGCTCATGCTCAAGAGAGATATTACAGGTATATCCAGCAGCTCCACTT 1147
 QY 1100 TTGATTCGAGGAGGTGTAGAAAAACAACAAGAAATGATGATGATGATGATGATGATGATG 1159
 DB 1148 TTGCTCCCAAGAGAGACATTTGCTCATGTCAAGATCCAGGCTACGATATTCCTGCTAAA 1207
 QY 1160 ACTCGAATTTGTTGAATGCTTGGCGATAGGAAGAGATCCCTAATCTTGGTGGGAACTT 1219
 DB 1208 ACAGAGATCATGATCAACGCTTACCGGATTCGACGTGATCCAAACTATGGAACAAACCTT 1267
 QY 1220 GGAAGTTTAAACCCAGAAAGGTTTAAAGATTGTGCAATTTGATTTATTAAGGGAGCGACATTT 1279
 DB 1268 GATGAGTTTAAACCCCTGATAGTTCTTGCAGCTCCATAGATTAACAGAGGACTGAACTTT 1327
 QY 1280 GAACTGGTACCATTTTGGTGCAGAAAAAGAAATATCTCTGGCATTACTTTCAGCTATTAC 1339
 DB 1328 GAGCTATTACCGTTTGGATCTGGTAGGAGATATGTCCAGGGATGACAAATGGGGATCGCC 1387
 QY 1340 AATTGGAGTATGTCATTATAAATCTATTATATCAATTTTAAATGGGAACCTGGCCGA 1395
 DB 1388 ATTGTTGAATGGGACTATTGAAATTTGCTTACTTCTCGACTGGGGGTACAGAA 1443

RESULT 9

US-10-953-349-13112
 ; Sequence 13112, Application US/10953349
 ; Publication No. US20060107345A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ALEXANDROV, Nikolai et al.
 ; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
 ; TITLE OF INVENTION: ENCODED THERBY


```
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 13112
; LENGTH: 1659
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (474)..(474)
; OTHER INFORMATION: n is a, c, g, or t
US-10-953-349-13112
```

```
Query Match          9.2%; Score 159; DB 6; Length 1659;
Best Local Similarity 56.4%; Pred. No. 5e-33;
Matches 341; Conservative 0; Mismatches 255; Indels 9; Gaps 2;

QY 807 TGGATGTTCTATTGTAATCTTCAAAAAAATGGAACGTTCCAGTGCAGTGACAAACGAAA 866
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
852 TTGATGTTTGTTCACATATCTGAAGATGAGAGTTCTGMAATTAAATTAACAAAGAAA 911
QY 867 GCATCAAGCATCCGTTTTCGAATGTTTCTGCCGGAGCGAACAACCTTGGAAAGCTA 926
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
912 ACAATTAAAGCCCTCATCTTGGACATTAATTTGCTGGGACTGACACCTCAGCTGTAA 971
QY 927 CAGAATGGTATGTCAGAGCTGATGAAAAATCCAACTGMACTAAGAAAAACACAAGAG 986
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
972 TGAATGGGCTATGGCAGAGTTAATCAACATCCAGGTGTGTGGAGAGGCGACAG 1031
QY 987 AAGTTAGACAAGTATTTGGTGAATGGGAAAGTTGATGAATCAAGATTTTCATGATTTGA 1046
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1032 NAATGGATGACGTGTTGGAAAGAGTAGAATAGTAGAAGATCAGATATTGCCAACTTC 1091
QY 1047 AATCTTCAAGTTAGTGTAAAGAACTCTAAGATTACATCTCGGTTGCTTGTATTC 1106
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1092 CTTACTTGCAGGCCATTTGTAGAGAAACATTAAGACCTTCACCAGCTGGTCCATTTGCTTT 1151
QY 1107 CGAGGGAGTGTAGAGAAACACACAGAAATTTGATGGATATGAATTCATCCGAACATCGAA 1166
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1152 TTAGAGAGTCATCTAGAGAGCTGTGGTTTGTGGGTATGATATTCAGCAAGACTCGAT 1211
QY 1167 TTGTTGTGAATCTTGGCGATAGGAAGAGATCTTAATCTTTGGTCGGAACCTGGAAGT 1226
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1212 TATTTGTCAATTTTGGGCTATTGGTAGGACCCCAATCACTGGGAGAACCTCTTTGAGT 1271
QY 1227 TTAAACCAGAGGTTTAAAGATTGTG-----CAATTGATTATAAGGGACGACATT 1278
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1272 TTAGCCAGAGAGGTTTGTGAAAATGGGAAGAGTCAATTGGATGTTAGGGGACACATT 1331
QY 1279 -TGAACCTGTACATTGTGTGAGGAAAAGAAATATGCTCGGCATTACTTCAGCTATTA 1337
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1332 ATCATCTACTTCGTTCCGTTAGTAGGAGAGAGCATGCTCTGTACTTCTTTGGCATTCG 1391
QY 1338 CCAATTGGAGTATGCTATTAATAAATCTATTATATCAATTTTAATTGGGAACCTGGCCGATG 1397
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1392 AAGTTGTGCATGTGAATTTGGCAGTCTTAATTCAGTGTTCATTCGAATGGAAGGTTGACTGTG 1451
QY 1398 GAATT 1402
Db      |||||
1452 ACAAT 1456
```

```
RESULT 10
US-10-953-349-17280
; Sequence 17280, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERY
; FILE REFERENCE: 2750-1579PUS2
```

```
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 17280
; LENGTH: 1658
; TYPE: DNA
; ORGANISM: Glycine max
; US-10-953-349-17280

Query Match          9.1%; Score 158.4; DB 6; Length 1658;
Best Local Similarity 49.6%; Pred. No. 7.2e-33;
Matches 643; Conservative 0; Mismatches 596; Indels 57; Gaps 7;

QY 164 CTCACCTTCTGATCTAGGCCATGAACGTTTATAGACCTTGGCTCAAAATTTATGGACCTGTT 223
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
164 CTTCTTCTCAACACACCTCACCAAGATTTTCAAGCTCTCACTCCGCTATATGGACCCATA 223
QY 224 ATGATCTTCAAAATGGCCAAAGTTTCAGCTGTTCTTCATTTCTTCAGCTGAAGCAGCCAAA 283
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
224 ATACACCTTTTCTTGTTCAGTCCCTGTGTGTGGCTTCCACAGCAAGAGCCGCCAAA 283
QY 284 GAGGTTATGAAAACCTCAGGCTGATGCCCTTCGCCCAACGCC---CTATCGTCTTTGGACGCA 340
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
284 GAGTTCTCAAACTCAGAACCCGCCCTTCTCCAACCGCCCGCTAACACTGTGCGCGTC 343
QY 341 CAGATTGTGTTTTATAATCGGAAAGATGTCCTTGTGTTTGTTCATATGGAGATCACTGGAGG 400
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
344 GAAACCTTAACCTACGGTTTCCAAGACTTCTTGTTCGCACCCCTACGACCCCTATTGGAG 403
QY 401 CAGATGAAGAAATTTGGATACTTGAATTTCTGAGTGCCAAAAAGTTTCAATCTCCAGG 460
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
404 TTCATGAAGAACTCTGCACTGCCAATCTCTCGGTGGCCACATCTCTGGAACGATTTCTT 463
QY 461 TTAATCCGAGAGGAAGAAATGGAGATGCCATCACATTCCTCCGTTTCAAAAGCCGATCT 520
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
464 CCCGTGAGACAGCAAGAGACAAAAAATTCATAAACCTGTCTTCCAAAAGGGTATTTCT 523
QY 521 CCGG-----TCAATATTACAAGATCATTTATGGCATT 553
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
524 GGTGAGGCGGTGGATTTTGGGGGAGAGTTTCATACGCTCTCGAACCAATCTGTGCGAGA 583
QY 554 ATAATTTTCATCATGATGAAGACATCCGTTTGGTAAATTTGAAGCAAAAAAGA-----AAGA 607
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
584 ATGATCGTGTGAGTCAGACGAGTACTACTGAGACAGAAAGAAAGTTGAAGAGATGAGGAAG 643
QY 608 TTGCTGAGTGTTCGCGATGTCAGTCAATGAGGCGACGAGTTTTGGCACCCGACAGCGT 667
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
644 CTGGTGAAGGATGCCGAGAGCTCTCGGGGAAGTTTCAACATATCGGACTTCGTTTCGTTTC 703
QY 668 TTTCCGAGCTGGAATTAATCTTCACTATATCATTTGAGCTGAGTCAAAACCCAGGCGTTTG 727
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
704 TTGAAGCGCTTTGATTTGACGGGTTCAACAGAGGCTCGAGAGATTCGGGAGCTGCTTT 763
QY 728 CATCAGGAGATTGACGATAT---ACTTGAAGAGATTTCTTAATGAACACAAAGCCCAATAA 783
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
764 GACACCGTTTGGACAGAAATCATTAAGCAACGTGAAGAGAAAGAAAGAACAGATGA 823
QY 784 GCCTTTTGAAGCGGATAACTTAA-----TGGATGTTCTATTGAATCTTCAAAAAAAT 835
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
824 ACCGTTGGAACAGAGATTTAAGGATATGCTTGATGTTTGTGTTGACATATCTGAAGAT 883
QY 836 GGAACCGTTCCAGTCCAGTGACAAACGAAAGCATCAAAAGCATCCGTTTGTCAAAATGTTT 895
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
884 GAGAGTTCTGAAATTTAAATTAACAAAGAAAAACATTAAGGCCCTTCATCTTGGACATATTA 943
QY 896 ACTGCCGAGGAGCAAAACCACTTCGAAAGCTACAGAAATGGGTAAATGGCAGAGCTGATGAAA 955
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
944 ATTGCTGGGACTGACACCTCAGCTGTAAAGTGAATGGGTATGGGTATGGCAGAGTTAATCAAC 1003
QY 956 AATCCAACCTGAATTAAGAAAAACACAAAGAAAGTTAGACAAGTATTTGGTGAATGGGA 1015
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1004 AATCCAGGTGTGTTGGAGAGGCAAGACAAGAAATGGATGCAATGCTGTTGGAAAGAGTAGA 1063
```

```
QY 1016 AAAGTTGATGAATCAAGATTTTCATGATTTTGAAATTTCTCAAGTTAGTGGTTAAAGAACT 1075
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1064 ATAGTAGAAGATCAGATATTTGCCAACCTTCTTACTTGCACGCAATGTTAGAGAAACA 1123
QY 1076 CTAAGATTACATCTCCGGTGTCTTGAATTCGAGGAGTGTAGAGAAACACACGAAT 1135
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1124 TTAAGACTTCACCCAGCTGGTCCATTGCTTTTATAGAGAGTCACTAGAGAGCTGTGGTT 1183
QY 1136 GATGGATATGAATTCATCCGACACTCGAATTTGTTGTAATGTTGGCGGATAGGAAGA 1195
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1184 TGTGGTATGATATTCAGCAAGACTCGAATTTTGTCAATGTTTGGGCTATTTGTAGG 1243
QY 1196 GATCCTAATCTTGGTCGGAACCTCGGAAGTTTTAAACCCAGAAAGTTTAAAGATTCTG-- 1253
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1244 GACCCCAATCACTGGGAGAACCTCTTGTAGTTTAGCCAGAGAGTTTGTGAAATGGG 1303
QY 1254 -----CAATTGATTATAAGGACGACATTT-GAACTGGTACATTTGGTGCAGGAAA 1306
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1304 AAGAGTCAATTGGATGTTAGGGGACAACTTATCATCTACTTCGGTTCGGTAGTGAAGA 1363
QY 1307 AGAATATGCTGGCATTTACTTCAGCTATTACCAATTTGGAGTATGTCATTATAATCTA 1366
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1364 AGAGCATGCTGGTACTTCTTTGGCATTTGCAAGTTGTGCATGTGAATTTGGCAGTTCTA 1423
QY 1367 TTATATCATTTTAAATTTGGGAACCTGGCGGATGGAATT 1402
Db 1424 ATTCAGTGTTCCAATGGAGGTTGACTGTGCAAT 1459
```

```
RESULT 11
US-10-953-349-10369
; Sequence 10369, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 10369
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-10369
```

```
Query Match      8.9%; Score 153.8; DB 6; Length 1722;
Best Local Similarity 53.2%; Pred. No. 1.2e-31;
Matches 326; Conservative 0; Mismatches 287; Indels 0; Gaps 0;
```

```
QY 853 AOTGACAAACGAAAGCATCAAGCATCCGTTTTCGAAATGTTTACTCCGGGAGCGAAAC 912
Db 953 ATTAACAATGAATGATCTCAACACATCTGCTTCGATGTGTTTGTGTCGCACACACAC 1012
QY 913 RACTTCGAAAGCTACAGATGGTAAATGGCAGAGCTGATGAAAATCCAACTGAACATAAG 972
Db 1013 AAACCTCTAGTACAATGGAAATGGCAATGACCGAGTTATTTCCGTAGCAGGAAAGATGGT 1072
QY 973 AAAAGCACAAAGAAAGTTAGACAAGTATTTGGTGAATGGGAAAGTTGATGAATCAAG 1032
Db 1073 CAAGACACAGATGAGATACGGCAAGTATTTGGTCAAAACGGTTCTGTTCAAGATCTGA 1132
QY 1033 ATTTCATGATTTGAAATCTTCAAGTTAGTGGTTAAAGAAACTTAAGATTACATCCTCC 1092
Db 1133 TATCCGAGTCTGCCTTACTTCAAGCAATTTGAAAGAGACTCTTCGTTTGCATCCTCG 1192
QY 1093 GGTGTGCTTGAATCCGAGGAGTGTAGAGAAACACACGAATTTGATGATATGAAATCA 1152
Db 1193 AGCTCCTTTGATCCCTTAGAAAAATCAGAAATCCGATGTTTCAGATTATGGGGTTCTCCTGTTCC 1252
```

```
QY 1153 TCCGAACACTCGAATGTTGTTGAATGCTTTGGCGGATAGGAAGAGATCTCTAATACTTGGTC 1212
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1253 TATAACACCCAGGTTGTTGTGAAACGTATGGGCGATAGACGAGACGCGAGCGTGTGGGA 1312
QY 1213 GGAACCTCGAAAGTTTAAACCCAGAAAGGTTTAAAGATTGTGCAATTTGATTTATAAGGAC 1272
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1313 AAATCCAATGAAGTTTCGAGCCAGAGAGGTTCTTGTGTACGAGAAACCGATGTAAGGCGAG 1372
QY 1273 GACATTTGAACTGGGTACCAATTTGGTCAGGAAAAAGAAATATGTCCTGGCATTTACTTCAGC 1332
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1373 AGATTTTCGAGTTGATACCGTTTGGATCAGGAAAGGATGTGTCCGGAAATCTCGATGGC 1432
QY 1333 TATTACCAATTTGGAGTATGTCAATATAATCTATTATATCAATTTTAAATGGGAACCTGGC 1392
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1433 TCTTAAGACAATGCATATGTGTCTTGCCTCTCTCTATTCTTTGACTGGAGCTTCA 1492
QY 1393 CGATGGAATTAACCTCAACACTTGTATGATGAACTGAAAGCTATTGGCGGTGCTCTCAGGAA 1452
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1493 AAACGGTGTGTCCTCCCGGAAACATTCACATGACGAGACTTTCGGTCTTACCTTACACAA 1552
QY 1453 AAAAATAGATCTT 1465
Db 1553 GGCCAAATCTCTT 1565
```

RESULT 12

```
US-10-953-349-21323
; Sequence 21323, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 21323
; LENGTH: 1660
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-21323
```

```
Query Match      8.8%; Score 152.2; DB 6; Length 1660;
Best Local Similarity 54.8%; Pred. No. 3.2e-31;
Matches 345; Conservative 0; Mismatches 278; Indels 6; Gaps 2;
```

```
QY 764 AATGAACACAAAGCCAAATAAGCCTTTTGAAGCGGATAACTTAATGGATGTTCTTATTGAAT 823
Db 823 AGTGGTAAAAGATATAATGACAAACAAGAGTCAAAGATATTATTGATATTCTTCTCCAG 882
QY 824 CTTCAAAAAAATGGAACGTTCCAGTGCAGTGACAAACGAAAGCATCAAGCATCCGTT 883
Db 883 CTACTTGTATGATCGTTTCTTCACTTTGATCTCACTCTCGACCACATAAAAGCCGTGCTC 942
QY 884 TTGCAAAATGTTTACTCCCGGAGCGAAACAACCTCGAAGCTACAGAATGGGTAAATGGCA 943
Db 943 ATGAACATCTTTTATAGCAGGAACAGACCCGAGTTCGCGACAATAGTTTGGGCAATGAAT 1002
QY 944 GAGCTGATGAAAAATCCCACTGAACTAAGAAAAAGCACAAAGAAAGTTAGACAAGTATTT 1003
Db 1003 GCACCTGTTGAAGAAATCCCAATGTGATGAGCAAGTTTCAAGGAGAGTGAAGAAATCTATTTC 1062
QY 1004 GGTGAAATGGGAAAAAGTTGATGAATCAAGATTTTCATGATTTGAAATTTCTTCAAGTTAGTG 1063
Db 1063 GGTGACAAAGATTTTATAAACCGAAGATGATGTGAAAGCCTTCTTTATCTCAAGACAGTG 1122
QY 1064 GTTAAAGAACTCTAAGATTA---CATCTCCGGTTGCTTCTGATTCGAGGAGGTGTAGA 1120
Db 1123 GTGAGGAGACATTAAGATTTATCCACCTTCACCACTACTTTTCCCAAGGGGTAACAATG 1182
QY 1121 GAAACACACGAATTTGATGGATATGAAATTTATCCGAACTATCGAATTTGTTGTGAATGCT 1180
```

Db	1183	GAACATGCAACATAGAGGGTACGAAATTC	AGCCAAAAC	TAGTGCATGTAATGCA	1242
Qy	1181	TGGCGCATAGGAAGAGATCCCTAAATAC	TCTGGTTCGGAAC	CTCGGAAGTTTAA	1240
Db	1243	TGGCCCATAGCAAGGACCCCTGAGAA	TTGGGAAGAGC	CTGAGAAATTTTCC	1302
Qy	1241	TTTAAACATGTGCAATTTGATATAAAGGG	A--CGACATTTGAAC	TGGTACCATTTGGT	1297
Db	1303	TTCTTCGAGATTCGATCGAGTTAAAGGG	GAATCATGAGTTA	AGGTGATCCCGTTGGT	1362
Qy	1298	GCAGAAAAGAATATGTCCTGGCATTA	CTTCACGTAATACCA	ATTTGGAGTATGTCATT	1357
Db	1363	TCTCGAAGGAGAAATGTGTCTCGCAAG	CAACATGGGAATTA	TGAATGTAGCTTTCTCTTG	1422
Qy	1358	ATAAATCTATTATATCATTTTTAAATGGGA			1386
Db	1423	CTAATCTCATTCACACGTTTGATTTGGAA			1451

```

RESULT 13
US-10-953-349-32482
; Sequence 32482, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 32482
; LENGTH: 1632
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-953-349-32482

```

551	Db	AACTGAGCGAGCGCATATCGCGGTACGTCCGGATCTTACGGTGGCGGCCATCGTCGGC	610
587	Qy	A- - - ATGTTAAGCAAAAAGAAAGATTGCTGAGTGTGGCGATGCGATCAATAGGCGAGCG	643
611	Db	AGCGGTTCAAGCAGAGGAGACAGTACCTGGAGATGCTGCAAGAGGAGCATCAAAATCGGTG	670
644	Qy	ACGAGTTTTGGCACCCGACGAGCGCTTTTCCGACGTGGAAATTACTTCACATATCATTTGGA	703
671	Db	CCCGGATGACCTTCCCGACATTTTCCCTCTCTCGGCTCTGTGGGCTCCTTAGCAGC	730
704	Qy	GCTGAGTCAAAACCCAGCGCTTTTGATCAGGAGATTGACGATATCTTTGAAGAGATTCTT	763
731	Db	GTCCCCGGCAGGATACAGCGCCATAGCCAAGACATGAAGCTGTTCTTAGACACCATCATC	790
764	Qy	ANTGAACACAAAGCAATAAGCCTTTTG-----AAGGGATACCTTAATG	808
791	Db	CAAGAGCACCAGGTGAACACGCGCCCGACTGCGACGGCGACAAAGAAGAAGACTTGTCTC	850
809	Qy	GATGTTCTATTGAACTTTCAAAAATAATGGAACGTTTCCAGTGGCCAGTGAACAAACGAAAGC	868
851	Db	GAGTGTCTTGAGACTTCAAAGGAGGGACTCGCAGTATCCACTCACCACTGACAAC	910
869	Qy	ATCAAGACATCCGTTTTCGAAATGTTTATCTGCCGGAGCGAAACAACCTTCGAAAGCTACA	928
911	Db	ATCAAGACCCGTCACTGCTGGACATGTTTGGCGCCGCGACGAGACGTCGGCGACGACGCTG	970
929	Qy	GAATGGGTAAATGCGAGAGCTGATGAATAATCAACTGAACTAAGAAAGACACAAGAAGAA	988
971	Db	CAGTGGCGGATGGCGAGCTGATACCGNACCCCGGGTTATGCGGAAGGCGCAAGACGAG	1030
989	Qy	GTTAGACAAGTATTTGGTGAATTTGGGAAATGTGATGAATCAAGATTTTCATGATTTGAAA	1048
1031	Db	GTCCGGCAGCAACTCGCGGGCAGCACAGGTGAGGAGGCCAGCTGGCAGATCTCGCA	1090
1049	Qy	TTCTTCAAGTTAGTGGTTAAAGAACTCTAAGATTACATCTCCGG--TTGCTTTGATT	1105
1091	Db	TACCTTGGGTTGTTCATCAAGGAGAGCGTGAGGATGCACCCGCGCGCCATTGCTGCTG	1150
1106	Qy	CCGAGGAGGTGTAGAGAAAACAACACGAATTTGATGGATATGAATTCATCCGAACACTCGA	1165
1151	Db	CCGCTAGGTGCGGAGCCCGTGTCAAGTTCTGGCCCTGACGTGCGGAGGGGTCATG	1210
1166	Qy	ATTGTTGTGAATGCTTGGGGGATAGGAAGATCTTAATACTTGGTCGGAACCTGGAAG	1225
1211	Db	GTGATCTGTAACCGTGGCGGCGATCGGCATCGAACCCGCGCACTGGGAGGACCCGAGGAG	1270
1226	Qy	TTTAAACCCAGAAAGTTTAAAGATTGTGCAATTTGATTAAGAGGACGACATTTGAACTG	1285
1271	Db	TTCCGACCGAGCGGTTTCGAGCAGAACCGGAGGACTTCAAGGGGCGGACTTCGAGTTC	1330
1286	Qy	GTACCATTTGGTCAGAAAAAGAAATATGTCCTGGCATTACTTCAGCTATTACCAATTG	1345
1331	Db	GTGCGGTTTCGCGCGCGGAGGAGATATGCCCGGCCATGGCTTTTCGGGCTTGGCGACGCTG	1390
1346	Qy	GAGTATGTCATTATAATCTATTATATCATTTTAAATGGGAACCTGGCCGATGGAATTACA	1405
1391	Db	GAGTCCGGCTCGCGCGCTGCTGTTCCACTTTCGACTTGGAGCTGCCAGCGGGCGCGCG	1450
1406	Qy	CCTCAAAACACTTGATATGACTGAAGCTATTGGCG	1439
1451	Db	GCCGAGGATCTGGACATGACCGAGGAGTTTGGCG	1484

RESULT 14
US-10-953-349-32468
; Sequence 32468, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349

; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 32468
; LENGTH: 1766
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-953-349-32468

Query Match 8.6%; Score 148.2; DB 6; Length 1766;
Best Local Similarity 47.5%; Pred. No. 3.7e-30;
Matches 614; Conservative 0; Mismatches 658; Indels 21; Gaps 5;

```
QY 114 CTCAGGGCCATGGAAGTTTCTTATCATAGTAATCTTCTCATTTTACTCACATCTCG 173
Db 183 CGCTGGCGGTGGCAGCTGCGGTGATCGGACCTGACACCTCGTCTGGCGGGC 242
QY 174 ATCTAGGCCATGAAGCTTTAGAGCTTGGCTCAAAATTTATGACCTGTTATAGTCTTC 233
Db 243 AGCTCCCTCACCGCGGATGCGGACCTGCGGCGAGCTACGGGCGGCGTCTGCTCC 302
QY 234 AATTTGGCCAGTTTCAGCTGTGTCTATTTCTTCACTGAAGCAGCCAAAGAGTTATGA 293
Db 303 AGCTGGCCAGGTGAAGACGCTGTGTCTCTCGGAGGGGCGCGGAGGTGATGA 362
QY 294 AAACCTCAGGCTGATGCTTTCGCCCAACGCCCTATCGTCTTGACGACACAGATTGTGTTT 353
Db 363 AGAACCCACACACCATGTTTGGCCACCGCGCGCTGAGCACCCATCGCGTCTCTCT 422
QY 354 ATAACTCGGAAGATGTCTGTTTGTTCATATGAGATCACTGGAGGAGATGAAGAAA 413
Db 423 ACGGGGCCAGGACATCGTCTTCGCCCTTACGGAGNACTTGGCGCCAGCTCCGCAAGA 482
QY 414 TTTGGATACTTGAATTTCTGAGTGCCAAAAGTTTCAATCTCCAGTTTAATCCGAGAG 473
Db 483 TCGCGTGTCCGAGCTCTTACCGCGCGCGGTCTCTCTCTTCCGCGCATCCGCGAG 542
QY 474 AAGAAATGGAGGATGCATCAATTTCTCTCG-----TTCGAAAGCCGGATCTCCGGTCA 527
Db 543 AGGAGTGCACCGCGCTCCGCGTCTCGGCGAGCTCGGCGGCGCGCGCGCTGG 602
QY 528 ATATTACAAAGATCATTTATGGCATTAATTTTCGATCATATGAAGAAATCCGTTGGTA 587
Db 603 AGATGCGCGGTGTCTATCCAGCTCGTACGAGACAGCAGCGCGCGCTCATAGGG 662
QY 588 A---TTGTAAAGAAAAGAAAGTTGCTGAGTGTTCGCCGATGCAGTCAATGAGGACGA 644
Db 663 ACCGGTGCAGGAGCGCGAGCTTCTCCGGAGCTCGACCGCATCGTGCAGCTCGGT 722
QY 645 CGAGTTTGGCACCGCAGACGCTTTTCCGACGTGGAATTAATCTTCACTATATCATTTGGAG 704
Db 723 CGGGGTTCAACTGGCCGACATGTGGCGCTCGTCGACGCTTGGCGGTGGCTCAGTGGAG 782
QY 705 CTGAGTCAAAACCCAGGCGTTTGCATCAGAGATTCAGCATATATCTTGAAGAGATTTTA 764
Db 783 CCGAGGAGTGGCCACACCTGTACACCATGCTCGACGGCATCGTTGAGGAGCACT-- 840
QY 765 ATGACACAAAGCCCAATAGCCTTTTGAAGCGGATTAATTAATGATGTTCTATTGAATC 824
Db 841 -GGAGAGGACGACGCGCGGAGGCGCACCGGAGACTTGTCTGACAGTGTCTGCTCAAGA 899
QY 825 TTCAAAATAAATGGAAACGTTCCAGTGCACGTGACAAACGAAAGCATCAAGCATCCGTTT 884
Db 900 TCCAGAGGAGGGTGGTCTCAGTTTCCCATCCACATGACGCGCTCAAGCCATCATCT 959
QY 885 TCGAAATGTTTATCGCGGAGCGGAAACAACTTTCGAAAGCTACAGAAATGGTAAATGGCAG 944
Db 960 TGGAGTATTCTCTGCGGCGAGTGAACAAACAACTACGACAATTGAGTGGGCGCATATCAG 1019
QY 945 ACCTCATGAATAATCCAACTGAACCTAAGAAAGACCAAGAAAGTTAGACAGTATTGG 1004
Db 1020 AGCTGATCAACAAACCGATGGCCATGACAGAGGCGACAGATGAGGTGCGACAGCCCTCC 1079
```

```
QY 1005 GTCAATGGGAAAAGTTGATGAATCAAGATTTTCATGATTTGAAATTTCTTCAAGTTAGTGG 1064
Db 1080 ATGCTAGTGAACCGGTGTCGAGCAGCCCTAAGCAGCTCCCATACCTCGCTTGTGTCA 1139
QY 1065 TTAAGAAACTCTTAAGATTACATCTCCGTTGTCTTGAATTCGAGAGGAGTGTAGAGAAA 1124
Db 1140 TCCGAGAGACGCTGCGGTACACCCGCCCTGCGGTGTTGTTCCGCGAGTGCAGGAAC 1199
QY 1125 CACACGAATTGATGGATGAAATTCATCCGAACTCGAATTTGTTGAATGCTTTGGG 1184
Db 1200 CGTGCCAAGTGCAGGATACGACGTGACGGGGCAGCAGGTTGGTCAATGCTTTGGG 1259
QY 1185 CGATAGAAAGAGATCCTAATACTTTGTCGGA---ACCTGGAAGTTTAAACCCAGAAAGGT 1241
Db 1260 CGTAGGCCGCGACGAGCGCTATTGGCCCGACGCGCTGAGGAGTTCCGGCCGAGACGGT 1319
QY 1242 TTAAGATTGTGA-----ATTGATTATAAAGGACGACATTTGAACCTGGTACCAATTG 1295
Db 1320 TCGAGGAAGAAGCAGCAAGGCGAGACTTTGGGGGTGCTGACTTTGGGTTCTTCCGTTTG 1379
QY 1296 GTGCGAGAAAGAAATATGCTCTGGCATTAATTCAGCTATTACCAATTTGGAGTATGTCA 1355
Db 1380 GCGCGCGCGCAGGATGTGCCCTGGGATGGCGTTGGCTCGCCGCGTCGAGCTCCCGC 1439
QY 1356 TTATAAATCTATTATATCATTTTAATTGGGAAC 1388
Db 1440 TTGCAAGCATGCTCTTCCACTTTGACTGGAGC 1472
```

RESULT 15

US-10-953-349-12509
; Sequence 12509, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 12509
; LENGTH: 1727
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (562)..(563)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1481)..(1481)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1484)..(1484)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1506)..(1506)
; OTHER INFORMATION: n is a, c, g, or t
US-10-953-349-12509

Query Match 8.5%; Score 147.2; DB 6; Length 1727;
Best Local Similarity 48.0%; Pred. No. 6.7e-30;
Matches 657; Conservative 0; Mismatches 675; Indels 36; Gaps 7;

```
QY 97 GAAGAAACAGATCCACCTCCAGGGCCATGGAAGTTTCTTATCATAGGTAATCTTCTCA 156
Db 122 GAATGCAAGAAATTAACCACTGGTCCAAAGGGTTGCCAATTTTGGGAAGCCCTTCACAA 181
QY 157 TTTATTACTCACTTCTGATCTAGGCCATGAACGTTTTAGAGCCTTTGGCTCAAATTTATGG 216
```

Db 182 GTTGGGCAAAATCCTCATCGTGACC-----TGCAAACTAGCCCCAAAATATGG 232
Qy 217 ACCTGTTATGAGTCTTCAAATGGCCAAAGTTTCAGCTGTGTGTCATTTCTTCAGCTGAAGC 276
Db 233 ACCGTCAATGACATACGTTTAGGTTTGTGCCGACCATAGTTGTTCTTCACCCAAATC 292
Qy 277 AGCCAAAGAGGTTATGAATACTCAGGCTGATGCCCTTCGCCCAAAGCCCTATCGTCTTGG 336
Db 293 TGCTGAATTTGTTCTCAAGACCCATGACCTTGTCTTTCTAGTAGACCAACGTTTCGTGGC 352
Qy 337 GCGCAGATTTGTTTATATCGAAAGATGCTCTTGTGTTCTTCATATGAGATCACTG 396
Db 353 GGATCAATACATCTCTGGGGCAGAGAACTTAGGCTTTGCTGAATATGTTCTTATTG 412
Qy 397 GAGGAGATGAAGAAATTTGATACCTTGAATTTCTGAGTGCCAAAAGTTCAATCCTC 456
Db 413 GCGCAACATGCGCAGATGTGCACATTTGGAATTTGTAAGCCAAATCCAAATTAATCCTT 472
Qy 457 CAGGTTAATCCGAGAGGAAGAAATGGAGGATGCCATCACAATTCCTCGTTCGAAAGCC-- 514
Db 473 CAGACGATGAGGAAGAGGAGCTTGACCTTTTGATCAAGCTTGTGAAGAGGAGGCCAA 532
Qy 515 ----GGATCTCGGTCAATATTACAAAGATCAATTTATGGCATTATAATTTGCATCATGAT 570
Db 533 TGATGGAGCTGCTGTTGATCTCAGTGTGCTCANNNGTTGCAACACTCAATTCAGACATGCTCTTG 592
Qy 571 AAGAACATCCGTTGG--TAAATTGTAAGCAAAAGAAAGATTTGCTGAGTGTTCGCGATGC 627
Db 593 TAGAATGATTTAGGAAGAGATGATGACCAAGACATGTTGGGAGAGGGTTCAAGGC 652
Qy 628 AGTCAATCAGGCGAGCGAGTTTGGCACC-----GCAGACGCTTTTCCGACGTGGA 680
Db 653 TGTGATTCAAGAAGCAATGCGTTTATTAGCAACTCCTTAACGTTAGGAGATTACATTCCTTA 712
Qy 681 AATTACTTCACTATATCATTTGAGAGCTGAGTCAAAACC--CAGGCGTTTGCATCAGGAGAT 738
Db 713 CATTTGGTCAATTTGACCTTCAAGGCGCTTAAAGCGCTTCAAGGTACTTTACGAAATCTT 772
Qy 739 TGACGATATCTTGAAGAGATTTCTTAATGAACACAAAGCAATTAAGCCTTTTGAAGCGGA 798
Db 773 CGATGACTTCTTCGAGAAAGTTATTGATGAGCACAATGGAATCAGAGAAGGGAGAGACAA 832
Qy 799 TA-----ACTTAATGGATGTTCTATTGAATCTTCAAAAAAATGGAAAACGTTCCAGTGCC 852
Db 833 GACCAAGGATTTTGTGGATGTCTGTTGGCTTTTGTAGTACTGAAGAAATCTGAATACCG 892
Qy 853 AGTGACAAACGAAAGCATCAAGCATCCGTTTTCGCAATGTTTACTCGCGGAGCGAAAC 912
Db 893 CATTTGAACGGTCCAAATATCAAGCTATATTGCTGGATAGTTGGCTGTTCAATGGATAC 952
Qy 913 AACTTCGAAAGCTACAGAAATGGGTAAATGGCAGAGCTGATGAAAAATCCAACTGAACTAAG 972
Db 953 TTCAGTACAGCAATTTGAGTGGACACTTTCAGAGCTACTAAAAATCCAAAGGTGATGAA 1012
Qy 973 AAAAGCACAAGAAAGTTAGACAAAGTATTGCTGAAATGGGAAAAGTTGATGAATCAAG 1032
Db 1013 GAAACTCCAAATGGAAATTTGGAACACTGTGTGGGTATGAAGAGGAAGTTGGGGGAATCGGA 1072
Qy 1033 ATTTTCATGATTTGAAATCTTCAAGTTAGTGGTTAAAGAACTCTAAGATTACATCCTCC 1092
Db 1073 CCTGGACAAGTTGAAAGTATTTGGAGATGGTTGTAAAGGAAAGCATGAGGCTCCATCCAGT 1132
Qy 1093 GGTTCGCTTG---ATTTCGAGGGAGTGTAGAGAAACACACGAATTTGATGGATATGAAAT 1149
Db 1133 GGTACCATTTGTAATACCAACCAATCCACGGAAGATTGCAATGTCGAGAGATTTTTCAT 1192
Qy 1150 TCATCCGAACACTCGAAATGTTTGAATGCTTTGGCGGATAGGAAGAGATCCCTAATCTTG 1209
Db 1193 ACCCAAAAAATCAAGGGTCATAATAAATGATGGGCAATTTATGAGAGACCCAAAGTGCATG 1252
Qy 1210 GTCGAAACCTGGAAGATTTAACCCAGAAAGGTTTAAAGATTGTGCAATTTATTATAAGG 1269

Db 1253 GGTGAGCGCAGAGAAGTTTCTGGCCAGAGAGATTTGAGGAAGCAACATAGATGTGAGAGG 1312
Qy 1270 GACGACATTTGAACTGGGTACCAATTTGGTGCAGGAAAAAGAAATATGTCTCTGGCAATTCTTC 1329
Db 1313 GCGTGACTTTCGAGCTGATACCAATTTGGGTCTGGCAGAAGGGCCTGTCCAGGATTGCAATT 1372
Qy 1330 AGCTATTACCNAATTTGGAGTATGTCAATTATAAATCTATTATATCATTTTAAATTTGGAACT 1389
Db 1373 GGGCCTAATAACGGTTCGTCAACAGTGGGCCAACTTGTGCAATGCTTTGATTGGAACT 1432
Qy 1390 GGGCGATGGAATTACACCTCAACACTTTGATATGACTGAAGCTATTGG 1437
Db 1433 GCCCAATAACATGTTTCCAGATGATTTGGACATGACAGAGCGGTTCCG 1480

Search completed: May 31, 2006, 10:17:50
Job time : 43 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 31, 2006, 09:04:09 ; Search time 351 Seconds
(without alignments)
9238.267 Million cell updates/sec

Title: US-10-759-813-1

Perfect score: 1733
Sequence: 1 gcaataaaggaaatggagc.....tcttttcaatccgaaaaa 1733

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_NA:*

- 1: /EMC Celerra_SIDS3/ptodata/2/ina/1 COMB.seq:*
- 2: /EMC Celerra_SIDS3/ptodata/2/ina/5 COMB.seq:*
- 3: /EMC Celerra_SIDS3/ptodata/2/ina/6A COMB.seq:*
- 4: /EMC Celerra_SIDS3/ptodata/2/ina/6B COMB.seq:*
- 5: /EMC Celerra_SIDS3/ptodata/2/ina/7 COMB.seq:*
- 6: /EMC Celerra_SIDS3/ptodata/2/ina/H COMB.seq:*
- 7: /EMC Celerra_SIDS3/ptodata/2/ina/PP COMB.seq:*
- 8: /EMC Celerra_SIDS3/ptodata/2/ina/RE COMB.seq:*
- 9: /EMC Celerra_SIDS3/ptodata/2/ina/RE COMB.seq:*
- 10: /EMC Celerra_SIDS3/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	399.4	23.0	1781	US-09-499-302A-1
2	369.8	21.3	1665	US-08-313-075A-29
3	367.4	21.2	1666	US-07-912-900-24
4	367.4	21.2	1666	US-08-285-309-24
5	367.4	21.2	1666	US-08-502-046-24
6	308	17.8	1691	US-08-948-564-3
7	253.2	14.6	1515	US-09-292-768-5
8	245.4	14.2	1762	US-08-881-784-5
9	245.4	14.2	1762	US-09-292-768-1
10	245.4	14.2	1762	US-09-292-768-65
11	243.8	14.1	1665	US-09-292-768-67
12	243.8	14.1	1762	US-09-292-768-63
13	242.2	14.0	1665	US-08-881-784-8
14	242.2	14.0	1665	US-09-292-768-3
15	242.2	14.0	1665	US-09-292-768-69
16	234.2	13.5	1762	US-09-172-339-5
17	224.2	12.9	1838	US-08-948-564-1
18	169	9.8	1488	US-09-689-783A-1
19	164.2	9.5	1812	US-08-313-075A-37
20	164.2	9.5	1824	US-08-606-505B-1
21	164.2	9.5	1824	US-09-616-990-1
22	162.6	9.4	1657	US-08-948-564-11
23	162.4	9.4	1755	US-07-912-900-29

24	162.4	9.4	1755	2	US-08-285-309-29	Sequence 29, Appl
25	162.4	9.4	1755	2	US-08-502-046-29	Sequence 29, Appl
26	161.4	9.3	1757	2	US-08-313-075A-49	Sequence 49, Appl
27	161.4	9.3	1757	3	US-09-142-108C-26	Sequence 26, Appl
28	159.6	9.2	1789	3	US-09-142-108C-1	Sequence 1, Appl
29	151.4	8.7	1539	3	US-10-142-231-53	Sequence 53, Appl
30	151.4	8.7	1539	5	US-10-884-115-53	Sequence 53, Appl
31	148.4	8.6	1722	3	US-09-033-055A-1	Sequence 1, Appl
32	146.6	8.5	1812	2	US-07-912-900-28	Sequence 28, Appl
33	146.6	8.5	1812	2	US-08-285-309-28	Sequence 28, Appl
34	146.6	8.5	1812	2	US-08-502-046-28	Sequence 28, Appl
35	142	8.2	1545	3	US-10-142-231-55	Sequence 55, Appl
36	142	8.2	1545	5	US-10-884-115-55	Sequence 55, Appl
37	139.6	8.1	1764	3	US-09-947-027-3	Sequence 3, Appl
38	137	7.9	1634	3	US-09-126-420A-2	Sequence 2, Appl
39	137	7.9	1737	3	US-09-126-420A-1	Sequence 1, Appl
40	133.2	7.7	1724	3	US-09-672-785-1	Sequence 1, Appl
41	127.2	7.3	1835	3	US-09-564-808-5	Sequence 5, Appl
42	127.2	7.3	1884	3	US-09-564-808-3	Sequence 3, Appl
43	126.4	7.3	1667	3	US-09-142-108C-22	Sequence 22, Appl
44	125.6	7.2	1838	2	US-09-091-432-1	Sequence 1, Appl
45	125.6	7.2	1838	3	US-09-387-663-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-499-302A-1
; Sequence 1, Application US/09499302A
; Patent No. 6369212
; GENERAL INFORMATION:
; APPLICANT: BOUNG-JUN, OH
; APPLICANT: MOON, KYUNG KO
; APPLICANT: YOUNG, SOON KIM
; TITLE OF INVENTION: A CYTOCHROME P450 GENE HIGHLY EXPRESSED IN THE
; TITLE OF INVENTION: INCOMPATIBLE INTERACTION
; FILE REFERENCE: 10324/P64443USO
; CURRENT APPLICATION NUMBER: US/09/499,302A
; CURRENT FILING DATE: 2000-02-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1781
; TYPE: DNA
; ORGANISM: Capsicum annuum
US-09-499-302A-1

Query Match	23.0%	Score 399.4;	DB 3;	Length 1781;
Best Local Similarity	56.3%	Pred. No. 8.6e-98;		
Matches	826;	Conservative	0;	Mismatches 611; Indels 30; Gaps 3;
QY	99	AGAAAAGAGATCCACCTCCAGGGCCATGGAAGTTTCTCTATCATAGTAAATCTTCTCATTT	158	
Db	86	AAAGCTAAACTTACCTCTCTGCTCCATGGAATACCTTTTATTGGAAGCCATACACT	145	
QY	159	TATTACTCATTCTCTATCTAGCCATGAACGTTTTAGAGCCCTGGCTCAAAATTTATGGAC	218	
Db	146	TGGCAGTGGCAGGTCCACTTCTCATATGGCCCTAAAAAATCTAGCAAAATTTATGGGC	205	
QY	219	CTGTTATGAGTCTTCAAAATTGGCCAAAGTTTTCAGCTGTTGTCTATTCTTCAGCTGAAGCAG	278	
Db	206	CGCTCATGCACTTACGCTCGGGGAAATTCCTACGTCATCAATTTGTCCTCCGCAATGG	265	
QY	279	CAAAAGAGGTTATGAAAACTCAGGCTGATGCTTCGCCCAACGCCCTATCGTCTTGGAGC	338	
Db	266	CGAAGGAAGTACTAAAAAATCAGACCTCGCTTTTCGCAACAGGGCGGAAACTTGTGGTG	325	
QY	339	CACAGATTGTTTATTAATCGGAAGATGCTGTTGTTCTCATATGAGATCACTGA	398	
Db	326	CTGACATCGTCCATTATGATAGTACGATATAGCAATTTCTCCATATGTTGAATACTGA	385	
QY	399	GSCAGATGAAGAAAATTTGGATACCTTGAATTTCTGAGTGCCCAAAAAGTTCAATCTCCA	458	

386	Db		GGCAGATT	CGTAAATTTG	CATCTCGAACT	CTCTTAGTGC	CAAGATGGT	CAAAATCTTTTA	445
459	Qy		GGTTAAAT	CCGAGAGGA	AAATGAGGAT	GCCATTC	CATTCCTCCG	TTCCAAAGCCGGAT	518
446	Db		GCTCAAT	TTCCGCGAG	GATGAGCTG	CGATGATG	TGTTCTCAT	CTATACGAAACCAT	505
519	Qy		CTCCGGT	CAATATTTAC	AAAAGATCAT	TTTATGGCAT	TTATAATTT	CGATCATGAT	578
506	Db		TTCCCGT	CAACCTTAC	AGACAAATA	TTTTGGTTT	TACAAGTT	CGGTAACTTTG	565
579	Qy		CCGTTGGTAA	--TTGTAAG	CAAAAAGAA	AGATGCTG	TAGTGTG	CCGATGCGAGT	635
566	Db		CTCTGGG	AAAAATAT	GTGTCGAC	CAAGACAA	CACTGATAA	TTTTTCATGAGG	625
636	Qy		AGGCAGCG	ACGAGTTTT	TGGCACC	CGACAGC	CTTTTCCG	ACGTGGAAT	695
626	Db		CAITGAC	AGGTGGAAT	TAGTATG	TGCTGAT	TTTTTTCC	TACATGGA	685
696	Qy		TCATTTG	GAGTCTCA	AAAAACCC	AGGCGTTT	TGCATCAG	GAGATGAC	755
686	Db		TTGGTGGTT	CAAAAAC	TAGACTG	CTGAAG	CGTCAAT	CGTAAAT	745
756	Qy		AGATTCTT	TAATGA	ACAAAGCC	CAATTAAG	CGCT-----	TTTG	791
746	Db		ATGTAGT	GATGATG	ACACAA	CAGAAATCG	AGCGGATG	CCAAAAAGG	805
792	Qy		AAGCGG	ATACCTTAA	TGGATGTT	CTATTGAAT	CTTCAAAAA	AAATGGAAC	851
806	Db		CGGTTGA	AGATTTG	CATCGATG	TTTTGCT	TAAAGGTT	TCGAGAAG	865
852	Qy		CAGTGACA	AAAGCAAA	AGCATCA	AAGCATCCG	TTTTTGCA	AAATGTTTAC	911
866	Db		CAATCA	CGGATGAC	AAATCA	ATATTAAGT	TGGACATG	TCTCCG	925
912	Qy		CAACTT	CGAAAGCT	CACAGAT	GGGTAAAT	TGGCAGAG	CTGATG	971
956	Db		CGTCAT	CGACAACT	ATATTAAT	TTGGGCAT	TAGCTGA	AAATGATCA	985
972	Qy		GAAAAG	CACAGAA	AGATGAG	ACAGTATTT	TGGTGA	AAATGCGG	1031
986	Db		CAAAAGGC	CAAGCTG	NAGTGAG	ACAAGCTTT	GAAAGG	AAAAAGAA	1045
1032	Qy		GATTTCA	TGATTTG	AAATTTCTT	CAAGTTAG	TGGTTTAA	AGAACTCTA	1091
1046	Db		ATCTTG	ATGAGTTG	NAGTACTTT	GAAATGAT	TAAATCA	AAAGAACT	1105
1092	Qy		CGGT---	TGCTTG	ATTCG	AGGAGTGT	TAGAGAAA	CAACAG	1148
1106	Db		CAATTC	CTCTAT	TAGTCC	TAGAGAA	TGATGA	AGGATAC	1165
1149	Qy		TTCATCC	GAACACT	CGAATTTG	TGTAATG	CTTGGCG	GATAGG	1208
1166	Db		TACCTTT	CAAAAC	CTCGAG	CTCATAG	TATTA	TGCAATGG	1225
1209	Qy		GGTCG	AACTCG	GGAAAGTTT	TAAACCC	CAGAAAG	GGTTTAA	1268
1226	Db		GGATG	ACCTTG	AAAGCTTTT	CCCC	CAGAGAT	TCGAGAT	1285
1269	Qy		GGACG	ACATTTT	GAACTGG	TACCATTT	TGGTGC	AGAAAA	1328
1286	Db		GAAGCC	ATCATCA	ATTTAT	TCCATTT	TGTCG	CGGAAG	1345
1329	Qy		CAGCTAT	TACCAAT	TTTGAG	TATGTC	ATATA	TATATCA	1388
1346	Db		TTGGTT	TAGC	CAATGTT	GGCAAC	CAACATT	TAGCTCA	1405
1389	Qy		TGGCCG	ATGGAAT	TACACT	CAAA	CACAT	TGATGACT	1448
1406	Db		TCCCT	TAATGG	ACAAAGT	CACG	AAATTTT	GGACATG	1465
1449	Qy		GGAAAA	AAAAAT	TAGATCT	TAA	GTGAT	TTCCTA	1508

Db	1466	CAAGAAAGGATGATCTCTGTTTGAATGCCACCCCTTATGATCCTTGAATGTAATTGAGACA	152
Qy	1509	TTTCTTTGATTACATAGAGGGTTGAAA	1535
Db	1526	GTTGTAGAAATAAAAAAGAGGGAGAAA	1552
RESULT 2			
US-08-313-075A-29			
Sequence 29, Application US/08313075A			
Patent No. 5639870			
GENERAL INFORMATION:			
APPLICANT: Holton, Timothy A.			
APPLICANT: Cornish, Edwina C.			
APPLICANT: Tanaka, Yoshikazu			
TITLE OF INVENTION: GENETIC SEQUENCES ENCODING FLAVONOID			
TITLE OF INVENTION: PATHWAY ENZYMES AND USES THEREFOR			
NUMBER OF SEQUENCES: 58			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Scully, Scott, Murphy & Presser			
STREET: 400 Garden City Plaza			
CITY: Garden City			
STATE: New York			
COUNTRY: U.S.A.			
ZIP: 11530			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: PatentIn Release #1.0, Version #1.25			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/08/313,075A			
FILING DATE: 30-NOV-1994			
CLASSIFICATION: 800			
PRIOR APPLICATION DATA:			
APPLICATION NUMBER: AU PL 1538/92			
FILING DATE: 27-MAR-1992			
PRIOR APPLICATION DATA:			
APPLICATION NUMBER: AU PL 6698/93			
FILING DATE: 07-JAN-1993			
PRIOR APPLICATION DATA:			
APPLICATION NUMBER: AU PCT/AU93/00127			
FILING DATE: 25-MAR-1993			
ATTORNEY/AGENT INFORMATION:			
NAME: DiGiglio, Frank S.			
REGISTRATION NUMBER: 31,346			
REFERENCE/DOCKET NUMBER: 9433			
TELECOMMUNICATION INFORMATION:			
TELEPHONE: (516) 742-4343			
TELEFAX: (516) 742-4366			
TELEX: 230 901 SANS UR			
INFORMATION FOR SEQ ID NO: 29:			
SEQUENCE CHARACTERISTICS:			
LENGTH: 1665 base pairs			
TYPE: nucleic acid			
STRANDEDNESS: single			
TOPOLOGY: linear			
MOLECULE TYPE: DNA			
FEATURE:			
NAME/KEY: CDS			
LOCATION: 3..1432			
US-08-313-075A-29			
Query Match 21.3%; Score 369.8; DB 2; Length 1665;			
Best Local Similarity 56.6%; Pred. No. 8.5e-90;			
Matches 803; Conservative 0; Mismatches 582; Indels 34; Gaps 5;			
Qy	95	TGGAAGAAACAGATCCACCTCCAGGGCCATGGAAGTTTCTCATATAGTAACTTCCT	154
Db	78	TGTCAACCAAAAAATTGCCTCCAGGCCCATGGAAGTACCTTTTCTTGGAAAGCTTGCT	137
Qy	155	CATTATTACTCACTTCGTATCTAGGCCATGAACGTTTATAGAGCCTTGGCTCAAAATTAT	214

Db 138 CATATGGTA-----GGTGGACTTCCACACCATGTCTCTTAGAGATTAGCCAAAAATAT 191
QY 215 GGACCTGTTATGAGTCTTCAAAATGGCCAAAGTTTCAGCTGTGTCTCAATTTCTTCAGCTGAA 274
Db 192 GGAACAAATATGACACTTCAACTAGTAAATTTCTGCCGTGTAGTTACTTCTCCTGAG 251
QY 275 GCAGCCAAAGAGGTTATGAAACTCAGGCTGATGCTTCGCCCAACGCCCTATCGTCTTG 334
Db 252 ATGGCAAGAAAGTACTAAAACTCATGACCTTGCATTTGCATATAGGCTTAACTTCTTA 311
QY 335 GAGGCACAGATTGTTTATATCGAAGATGCTTGTGCTTCATATGAGATCAC 394
Db 312 GCAATTGAGATTGCTGCTATATAGTTGAGCAATTCGCTTTTCCCGTATGCTGATTAC 371
QY 395 TGGAGCAGATGAAGAAAAATTTGGATACTTGAAATTTCTGAGTGCACAAAAAGTTCAATCC 454
Db 372 TGGAGCAATGGTAAATTTGTGTATTTGGAAGTCTTAGTGCACAAAAATGTCGGTCA 431
QY 455 TCCAGGTTAATCCGAGAGGAAGAAATGGAGATGCGATCACAATTCCTCCGTTTCGAAGCC 514
Db 432 TTTAACTCGATTAGACGAGATGAATACTCTTATGATCGATTTTTCGATCATCTTCT 491
QY 515 GSATCTCCGCTCAATATTAACAAGATCATTTATGCAATTAATTTTCGATCATGATAAGA 574
Db 492 GGTAAAGCCAGTTAATATAACAGAAAGGATCTTTTCAATTCACAAGCTCTATGATTTGTAGA 551
QY 575 ACATCCGTTGGTAATTG---TAAGCAAAAGAAAGATTGCTGAGTGTCCCGATGCAGTC 631
Db 552 TCAGTATTTGGGAAAGAAATAAAGGAAAGAGATGTATAGCATGTGAAAAAATG 611
QY 632 AATGAGGACGACGAGTTTGGCACCGCAGCGCTTTTCCGACGTGGAAATTTACTTTCAC 691
Db 612 ACAGGCTTAATAGATGGGTTCGATGTGGCTGACATATCCCTTCGTTGAGGTTTCTTCAT 671
QY 692 TATATCATTTGAGCTGAGTCAAAACCCAGCGCTTTCATCAGAGATTCACATATCTT 751
Db 672 GTACTAATCGGTATGAAGGGTAAAAATATGATGATGTTTCATGTAAGGTAGATCTATTGTT 731
QY 752 GAAGAGATTCTTAATGAACACAAAG-----CCAAATAGCCTTTT 790
Db 732 GAGGAAGTATGATGAGCACAAGAAACTCTTCGAACTGGCAAGCCAAATGGTGAAGTG 791
QY 791 GAAGCGGATTAATGATGTTCTTATGAAATCTTCAAAAAATTTGAAACGTTTCAGTG 850
Db 792 GGAGGAGAAATTAATGATGTTATGCTTAAGACTTAAGGAAGAGGAGACCTTCAACTT 851
QY 851 CCAGTGACAAACGAAGCATCAAGCATCCGTTTTCGAAATGTTTACTGCCGGGAGCGAA 910
Db 852 CCAATCACAAAATGACAAACATCAAGGCCATTTTAAATGACATGTTTGTCTGCGGGAACAGAA 911
QY 911 ACAACTTCGAAAGCTACAGATGGGTAAATGGCAGAGCTGATCAAAAAATCCAACTGAATTA 970
Db 912 ACTTCATCAACAATTAATCTGGGCCATGGTAGAACTGATGAAATAATCCAAGTGTATTC 971
QY 971 AGAAAGCACAAGAAAGATTTAGACAAGTATTTGGTGAATGGGAAAAAGTTGATGAATCA 1030
Db 972 GCGAAAGCTCAAGCAGAGAGTAAGAGAAGTCTTCAAGGGGAAGAAACTTTCGATGAAGAT 1031
QY 1031 AGATTCATGATTGAATTTCTCAAGTTAGTGGTTAAAGAACTCTAAGATTACATCT 1090
Db 1032 GATATCGAGGAGCTGAATTTACCTTAAGTTAGTCAATTAGAGAAACTTTAAGACTCCACCT 1091
QY 1091 CCGGTT---GTCTGATTCGAGGAGTGTAGAGAAACAAACGAATTTGATGATGA 1147
Db 1092 CCACTTCACCTTTGCTTCCAAGAGATGTGGAGAGAAACAGAAATAAATGGCTACACT 1151
QY 1148 ATTCATCCGAACACTCGAATTGTTGTGAATGCTTGGCGCATAGGAAGAGATCCTAATACT 1207
Db 1152 ATTCCTTTAAATACCAAGTCTAGTTAAATGTTTGGGCTATTGGAAGAGATCCAAATAT 1211
QY 1208 TGGTCGGAACCTGGAAAGTTTAAACCCAGAAAGGTTTAAAGATTGCGAATTGATTATAAA 1267
Db 1212 TGGGATGATGCAGAAAGCTTTAAGCCTGAGAGATTGAACATACTCTTTGAAATTTGCT 1271

QY 1268 GGGACGACATTTGAACCTGGTACCATTTGGTGAGGAAAGAAATATGTCTCGCATTA 1327
Db 1272 GCAATAATTTTGAATATCTCTCTTTTGGTAGTGGAAAGAGGATTTGCCCGGAATATCA 1331
QY 1328 TCAGCTATTACCAATTTGGAGTATGTCAATTATAAATCTATTATATCATTTTAATTGGGA 1387
Db 1332 TTTGGTTTAGCTAATGTTTATCATCCATTTGGCTCAATTTGTATCATTTTCGATTGAGA 1391
QY 1388 CTGGCCGATGGAATTACACTCAACACATTTGATATGACTGAAGCTATTGGCGGTGCTCTC 1447
Db 1392 CTTCTACTGGGGTGCAGCCAAATGACTTTTGAATTGACT-AGTTAGCTGGAGTAACTACT 1450
QY 1448 AGCAAAAAAATAGACTTAAAGTTGATTCCTATTCCATAT 1486
Db 1451 GGTAGAAAAGAGACCTTTACTTGATTTTCACTCCTTAT 1489

RESULT 3

US-07-912-900-24
; Sequence 24, Application US/07912900
; Patent No. 5349125
; GENERAL INFORMATION:
; APPLICANT: Holton, Timothy A.
; APPLICANT: Cornish, Edwin C.
; APPLICANT: Kovacic, Filipa
; APPLICANT: Tanaka, Yoshikazu
; APPLICANT: Lester, Diane R.
; TITLE OF INVENTION: GENETIC SEQUENCES ENCODING FLAVONOID
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/912,900
; FILING DATE: 19920713
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Digilio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8633
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1666 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-07-912-900-24

Query Match 21.2%; Score 367.4; DB 2; Length 1666;
; Best Local Similarity 55.7%; Pred. No. 3.8e-89;
Matches 824; Conservative 0; Mismatches 621; Indels 34; Gaps 5;
QY 35 TTTCCGACATTTAATAAGTTTCTGCTTGTGTTTAACTTAGTAGTAGTAGAGTTG 94
Db 19 TTTCTTCTCTCTTATGTTATTTTCCCTCATTTTTCATTAAGAAAATGAAGAAATCCAA 78
QY 95 TGAAGAAACAGAAATCCACCTCCAGGGCCATGGAAGTTTCTCTATCATAGTAATCTTCT 154

Db 79 TGTCAAAACCAAAATTTGCTCCAGGCCCATGAAAGTACCTTTCTTGGAGCTTGCTT 138
Qy 155 CATTTTATTACTCAGTCTGATCTAGGCGCATGAACGTTTTAGAGCCTTGCGTCAAAATTTAT 214
Db 139 CATATGGTA-----GGTGGACTTCCACACCATGTCTTAGAGATTTAGCCMAAAATAT 192
Qy 215 GGACCTGTTATGAGTCTTCAAAATTTGGCCAAAGTTTCAGCTGTGTGTCATTTCTTCAGCTGAA 274
Db 193 GGACCAATATGACCTTCAACTAGTGTAAATTTCTGCGGTGTAGTTACTTCTCCTGAG 252
Qy 275 GGAGCCAAAGAGGTATGAAGAACTCAGGCTGATGCTTGGCCCAAGCCCTATCGTCTTG 334
Db 253 ATGGCAAGAAAAGTAGTAAACTCATGACCTTGCATTTGTCATATAGGCCCTAAACTTTCTA 312
Qy 335 GAGGCACAGATTGTGTTTATTAATCGGAAAGATGCTTGTGTTTCATATGAGATCAC 394
Db 313 GGCATTTGAGATTGTCTGCTATTAATAGTTCAGACATTTGCCCTTTCCCGTATGTTGATAC 372
Qy 395 TGGAGGCAGATGAAGAAAATTTGGATACCTTGAATTTCTGAGTGCACAAAAGTTTCAATCC 454
Db 373 TGGAGGCAATCGGTAAATTTGTGATTTGGAAGTCTTAGTGCACAAAATGTCGGTCA 432
Qy 455 TCCAGGTTAATCCGAGAGAAATGGAAGATGCGCATCACATTCCTCGTTCGAAAGCC 514
Db 433 TTTAACTCGATTAGACGAGATGAAATACCTTCTTATGATCGATTTTTCGCGATCATCTCTC 492
Qy 515 GGAATCTCGGTCAATATTAACAAGATCATTTATGCAATATTAATTTTCGATCATGATAGA 574
Db 493 GGTAAAGCCAGTTAATAATAACAAGAGATCTTTTCAATCAAGCTCTATGATTTGTAGA 552
Qy 575 ACATCCGTTGGTAATG---TAAGCAAAAAGAAAGATTGCTGAGTGTGTCGATCAGTC 631
Db 553 TCAGTATTGGGAAAGATTAAGGAAAGCAAGATGATACGACATGTGAAAAAATG 612
Qy 632 ANTAGGCGACGAGATTTTGGCACCAGCAGAGCGCTTTCCAGCTGGAAATTTACTTCAC 691
Db 613 ACAGGCTTAATAGATGGGTTCGATGTGGCTGCATATTCCTTCGTGAGGTTTCTTCAT 672
Qy 692 TATATCATTTGAGCTGAGTCAAAACCCAGCGCTTTGTCATCAGGAGATTGACGATATCTT 751
Db 673 GTACTAATCGGTATGAGGGTAAATATGATGTTTCATCGTAAGGTAGATCTATTGTT 732
Qy 752 GAAGAGATTCTTAATGAACACAAAG-----CCAATAAGCCTTTT 790
Db 733 GAGGAAGCTATGAATGAGCACAAAGAAACTCTTCGAACTGGCAAGACCAATGGTGAAGTG 792
Qy 791 GAAGCGGATAACTTAATGATGTTCTATTGAATCTTCAAAAATGGAACCGTTCCAGTG 850
Db 793 GGAGGAGAAAGATTTAATGATGTTTGTCTAAGACTTAAAGGAGAGGAGACCTTCAACTT 852
Qy 851 CCAGTCACAAACGAAAGCATCAAGCATCCGTTTTCGAAATGTTTACTCCGCGGAGCGAA 910
Db 853 CCATCACAATATGACAACACTAAAGCCATTTTATGACATGTTTCTCGCGGACAGAA 912
Qy 911 ACAATCTCGAAAGCTACAGATGGTAAATGGCAGAGCTGATGAAAAATCCAACTGAACATA 970
Db 913 ACTTCAATCAACACAAATTAATGCGGCCATGGTAGAATGATGAAAAATCCAGTGATTC 972
Qy 971 AGAAAGCAAGAGAGATTAGACAAGTATTTGGTGAATGGGAAAGTTGATGATATCA 1030
Db 973 GCGNAAGCTCAAGCAGAGGTAAAGAGATCTTCAAAAGGGAAGAACTTTTCGATGAAGAT 1032
Qy 1031 AGATTTTCATGATTGAAATCTTCAAGTTAGTGTGTTTAAAGAACTCTAAGATTATACATCT 1090
Db 1033 GATATCGAGAGCTGAAATTAACCTTAAGTTAGTCAATTTAGAGAACTTTAAGATCTACCTT 1092
Qy 1091 CCGGTT---GTCTTGAATCCGAGGGAGTGTAGAGAAACACAGAAATGATGATGAA 1147
Db 1093 CCACCTTCCACTTTTGTCTTCAAGAGAAATGTCGAGAGAGAAACAGAAATAAATGCTACACT 1152
Qy 1148 ATTCATCCGAACTCGAATTTGTTGAATGCTTGGCGGATAGAGAGATCCTAATCT 1207

Db 1153 ATTCCTTTAAATACCAAGTCATAGTTAACTGTTTGGGCTATTGGAAGAGATCCAAATAT 1212
Qy 1208 TGGTCGGAACCTGGAAGTGTAAACCAAGAAAGTTTAAAGATTTGCAATTTGATTAATAA 1267
Db 1213 TGGGATGATGCAGAAAGCTTTTAAGCCTGAGAGATTTGAACATAACTCTTTTGAATTTTGT 1272
Qy 1268 GGGAGGACATTTGAATGCTGACCAATTTGGTGCAGGAAAGAAATATGCTTGGCATTTACT 1327
Db 1273 GGCAATAATTTTGAATATCTTCTCTTTTGGTAGTGGAAAGGAGATTGCCCCGGAATATCA 1332
Qy 1328 TCAGCTATTACCAATTTGGAGTATGTCAATATAAATCTATTATATATCATTTTAAATTGGGAA 1387
Db 1333 TTTGGTTTAGCTAATGTGATCATCCATTTGCTCAATTTGTTGATCAITTCGATTGGAGA 1392
Qy 1388 CTGCCCAGATGGAATTTACACCTCAACACTTGATATGATGAGTGAAGCTATTGGCGGTCTCTC 1447
Db 1393 CTTCCTACTGGGTGCGACCCAAATGACTTTGAAATGACT-AGTTAGCTGGAGTAACACTACT 1451
Qy 1448 AGGAAAAAATAGATCTTAAGTTGATTCCTATTCCATAT 1486
Db 1452 GGTAGGAAAAGAGACCTTTACTTGTATTTCACCTCTTAT 1490

RESULT 4

US-08-285-309-24
; Sequence 24, Application US/08285309
; Patent No. 5569832
; GENERAL INFORMATION:
; APPLICANT: Holton, Timothy A.
; APPLICANT: Cornish, Edwin C.
; APPLICANT: Kovacic, Filipa
; APPLICANT: Tanaka, Yoshikazu
; APPLICANT: Lester, Diane R.
; TITLE OF INVENTION: GENETIC SEQUENCES ENCODING A 3,5'-
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/285,309
; FILING DATE: 03-AUG-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8633Z
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1666 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-285-309-24

Query Match 21.2%; Score 367.4; DB 2; Length 1666;
Best Local Similarity 55.7%; Pred. No. 3.8e-89;
Matches 824; Conservative 0; Mismatches 621; Indels 34; Gaps 5;

QY 35 TTTCCGAGCATTTTAAAGTTTCTGCTGTTTAAATCTTAGTAGTCAATGAGGTTG 94
DB 19 TTTCCCTTTCTCCTTATGTTATTTTCTCCTCATTTTCAATTAAGAAATGGAAGAAATCCAAT 78
QY 95 TGAAGAAACAGAAATCCACTCCAGGCCATGGAAGTTTCTTATCATAGGTAATCTTCCCT 154
DB 79 TGTCAAAACCAAAATTTGCTCTCCAGGCCATGGAAGTACCTTTTCTTGGAAAGCTTGCTT 138
QY 155 CATTTATTAATCTACTTCTGATCTAGGCTAGGCAAGCTTTTAGAGCCTTTGGCTCAAAATTTAT 214
DB 139 CATATGGTA-----GGTGGACTTCCACACCATGTCTTAGAGATTTAGCCAAAATAT 192
QY 215 GGAACCTGTTATGAGTCTTCAAAATGCGCAAGTTTCAGCTGTTGTGCAATTTCTTCAGCTGAA 274
DB 193 GGAACCAATTAATGACCTTCAACTAGGTAAATTTCTGCGGTGTAGTTACTTCTCCTGAG 252
QY 275 GCAGCCAAAGAGTTATGAACCTCAGGCTAGTGCCTTCGCCCAAGCCCTATCGTCTTG 334
DB 253 ATGGCAAGAAAGTACTAATAAATCTATGACCTTTGCAATTTGCAATATAGGCTTAAACTTCTA 312
QY 335 GAGCGACAGATTGTTGTTTATATCGAAAGATGCTGTTTGTCTTCAATGAGATCAC 394
DB 313 GGAATGAGATTGCTGCTATATAATGTTTCAGACATTTGCCCTTTTCCCGTATGTTGATTAC 372
QY 395 TGGAGCAGATGAAGAAAATTTGGATACTTGAATTTCTGAGTGCACAAAAGTTCAATCC 454
DB 373 TGGAGGCAATGCGTAAATTTGTGATTGGAAGTCTTAGTGCCAAAATGTCGGTCA 432
QY 455 TCGAGTTAATCGGAGAGAAAGAAATGGAAGTGCATCACATTCCTCGCTTCGAAGCC 514
DB 433 TTAACTCGATTAGACGAGATGAATACTTCTTATGATCGATTTTTCGTCATCTCTC 492
QY 515 GGATCTCCGCTCAATATTAACAAGATCATTTATGGCATTATATTTTCGATCATGATAAGA 574
DB 493 GGTAAAGCAATTAATAACAAGAGATCTTTTCAATCACAGCTCTATGATTTGTAGA 552
QY 575 ACATCCGTTGGTAATG---TAAGCAAAAAGAAAGATTGCTGAGTGTTCGGATGCAGTC 631
DB 553 TCAGTATTTGGGAAAGAAATAAAGGAGAAAGCAAGATGATACGACATGTGAAGAAATG 612
QY 632 ATGAGGAGCAGAGTTTGGCAGCGCAGAGCTTTTCCGACGTGGAAATTAATCTTAC 691
DB 613 ACAGGCTTAATAGATGGGTTGATGTTGCTGACATATTCCTTCGTGAGGTTTCTTCA 672
QY 692 TATATCATTTGGAGCTGAGTCAAAACCCAGCGTTTTCATCAGAGATTCAGCATATCTT 751
DB 673 GTACTAATCGGTATGAGGTTAAATATGATGTTCAATCGTAAGTAGATGCTATTGTT 732
QY 752 GAAGAGATTTTAATGAACACAAAG-----CCAATAAGCCTTTT 790
DB 733 GAGGAAGCTATGAATGAGCACAAAGAAACTTTCGAACCTGGCAAGACCAATGGTGAAGTG 792
QY 791 GAAGGGGATACTTAATGATGTTCTTAATGAATCTTCAAAAATAAAGGAAACGTTCCAGTG 850
DB 793 GGAGGAGAAATTAATGATGTTATGCTTAAGACTTTAAGGAAGAGGAGACCTTCAACTT 852
QY 851 CCAGTGACAAACGAAGCATCAAGCATCCGTTTTCGCAATGTTTACTCCGCGGAGCGAA 910
DB 853 CCAATCACAAAATGACACACTAAAGCCATTTTAATGACATGTTTCTCGGGAAACAGAA 912
QY 911 ACAACTTCGAAGCTACAGAATGGGTAATGGCAGAGCTGATGAAAATAATCCAATCACTA 970
DB 913 ACTTCATCAACAACAAATTAACCTGGGCCATGGTAGAACTGATGAAAATCCACGTGATTTC 972
QY 971 AGAAAAGCACAGAAAGTTAGACAGTATTTTGGTGAATGGGAAATGGGAAAGTTGATGAATCA 1030
DB 973 GCGAAAGCTCAAGCAGAGGTAAGAGAAGTCTTCAAGGGGAAAGAACTTTTCGATGAAGAT 1032
QY 1031 AGATTTTCATGATTGAAATTTCTCAAGTTAGTGGTTTAAAGAACTCTAAGATTACATCT 1090
DB 1033 GATATCGAGGAGCTGAATACCTTAAGTTAGTCAATTAGAGAACTTTAAGATCTCACCT 1092
QY 1091 CCGGTT---GTCTTGATTCGAGGGAGTGTAGAGAAACACAGAAATTTGATGGATGAA 1147

DB 1093 CCACTTCCACTTTTGTCTTCCAGAGAAATCTCGAGAGAAACAGAAATAAATGCTACACT 1152
QY 1148 ATTCAATCCGAACACTCGAAATTTGTTGTAATGCTTGGGCGATAGGAAGAGATCTTAATACT 1207
DB 1153 ATTCTTTTAAATACCAAGCTCATAGTTAATGTTTGGGCTATTGGAAGAGATCCAAAATAT 1212
QY 1208 TGTGCGAACCTTGGAAAGTTTAAACCCAGAAAGTTTAAAGATTGTCAATTGATTATATAA 1267
DB 1213 TGGGATGATGCAAGAAAGCTTTAAGCCTGAGAGATTTGAACATAACTCTTTGAAATTTGCT 1272
QY 1268 GGAAGCAGACTTTGAACCTGGTACCATTGGTGCAGGAAAAAGAAATATGCTCGGCATTAAT 1327
DB 1273 GCAATAATTTTGAATATCTTCTTTTGGTAGTGAAGAGGATTTGCCCGCAATATCA 1332
QY 1328 TCAGCTATTACCAATTTGGAGTATGTCATATAAATCTATTATATCATTTTAATTTGGAA 1387
DB 1333 TTTGGTTTAGCTTAATGTGTATCATCCATTGGCTCAATTTGTTGATCATTTTCGATTGGAGA 1392
QY 1388 CTGGCGGATGGAATTACACCTCAAAACACTTTGATATGACTGAAGCTATTGCGGTGCTCTC 1447
DB 1393 CTTCTACTGGGTGACCCAAATGACTTTGAATTGACT-AGTTAGCTGGAGTAACACT 1451
QY 1448 AGGAAAAAATAGATCTTAAAGTTGATTTCTCTATTTCATAT 1486
DB 1452 GGTAGGAAAGAGACCTTTACTTTGATTTCACCTCTTAT 1490
RESULT 5
US-08-502-046-24
; Sequence 24, Application US/08502046
; Patent No. 5861487
; GENERAL INFORMATION:
; APPLICANT: Holton, Timothy A.
; APPLICANT: Cornish, Edwina C.
; APPLICANT: Kovacic, Filipa
; APPLICANT: Tanaka, Yoshikazu
; APPLICANT: Lester, Diane R.
; TITLE OF INVENTION: GENETIC SEQUENCES ENCODING A 3,5'-
; TITLE OF INVENTION: HYDROXYLASE AND USES
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESS: Scully, Scott, Murphy & Presser
; CITY: Garden City
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/502,046
; FILING DATE: 14-JUL-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/285,309
; FILING DATE: 03-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 86332
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1666 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

```
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-502-046-24

Query Match      21.2%; Score 367.4; DB 2; Length 1666;
Best Local Similarity 55.7%; Pred. No. 3.8e-89;
Matches 824; Conservative 0; Mismatches 621; Indels 34; Gaps 5;

QY 35 TTTCCGAGCATTTTAATAAGTTTCTGCTGTGTTTAAATCTTAGTAGTAGTCATGAGGTG 94
Db 19 TTTCTCTCTCTTATGTTATTTCCCTCAATTTTCAATTAAGAAATGGAAGAAATCCAAT 78
QY 95 TGAAGAAACAGAAATCCACTCCAGGCCATGGAAGTTTCTTATCATAGGTAAATCTTCT 154
Db 79 TGTCAAAACCAAAAAATGCTCCAGGCCATGGAAGTACCTTTTCTTGGAGCTTGCTT 138
QY 155 CATTTATTACTCACTTCTGATCTAGGCGCATGAACTTTTAGAGCCTTGCTCAAAATTTAT 214
Db 139 CATATGGTA-----GGTGGACTTCCACACCATGTCCTTAGAGATTAGCCAAAAATAT 192
QY 215 GGACCTGTTATGAGTCTTCAAAATGGCCAAAGTTTCAGCTGTGTGTCATTTCTTCAGCTGAA 274
Db 193 GGACCAATATGCACTTCACTAGGTAAATTTCTGCGTGTAGTTACTTCTCTGAG 252
QY 275 GCAGCCAAAGAGGTTAAGAAACTCAGGCTGATGCTTGGCCCAAGCCCTATCGTCTTG 334
Db 253 ATGGCAAGAAAAAGTACTAAAAACTCATGACCTTGCAATTTGCATATAGGCCCTAAACTTCTA 312
QY 335 GAGCACAGATGCTGTTTATATCGGAAGATGCTTGTGTTGCTTCATATGAGATCAC 394
Db 313 GGCATTGAGATGCTGCTATATATAGTTCAGACATTTGCCCTTTCCCGGTATGGTATAC 372
QY 395 TGGAGCAGATGAAGAAAAATTTGGATACCTTGAATTTCTGAGTGCACAAAAAAGTTCAATCC 454
Db 373 TGGAGCAATCGTAAATTTGTTATTTGGAAGTCTTTAGTGCACAAAAATGTCGGTCA 432
QY 455 TCCAGGTTAATCCGAGAGAAATGGAAGATGCCATCACATTTCTCGTTCGAAGCC 514
Db 433 TTTAACTCGATTAGACGAGATGAAATACTTCTTATGATCGATTTTTCGATCATCTCTC 492
QY 515 GGATCTCGGTCAATATACAAAGATCAATTTATGGCATATATATTTGCAATCATGATAGA 574
Db 493 GGTAGCCAGTTAATATAACAGAGGATCTTTTCAATCACAAGCTCATGATTTGTAGA 552
QY 575 ACATCGGTTTGGTAATTTG---TAAGCAAAAGAAAGATTGCTGAGTGTTCGCGATCGATC 631
Db 553 TCAGTATTTGGAAAGAAATAAAGGAGAAAGCAATGTATACGACATGTGAAAAAATG 612
QY 632 AATGAGGACGACGAGTGTTCGACCGCAGAGCGCTTTCCGACGTGGAATTAATCTTAC 691
Db 613 ACAGGCTTAATAGATGGGTTGATGTGCTGACATATTCCTTCGTTGAGGTTTCTTCAT 672
QY 692 TATATCATTTGAGCTGAGTCAAAACCCAGCGTTTTCATCAGAGATTCACGATATCTT 751
Db 673 GTACTAATCGGTATGAAGGTTAAATATGATGTTTCAATCGTAAGTAGATGCTATTGTT 732
QY 752 GAAGAGATTTCTTAATGAACACAAAG-----CCATAAGCCTTTT 790
Db 733 GAGGAGCTATGAATGAGCACAAGAAATCTTTGGAATGCGCAAGCAATGTTGAAGTG 792
QY 791 GAAGCGGATTAATGATGTTCTTATGATCTTCAAAATAAGAAATGGAACGTTCCAGTG 850
Db 793 GGAGGAGAAAGATTTAATGATGTTATGCTAAGACTTTAAGGAAGAGGAGACCTTCAACTT 852
QY 851 CCAGTCACAAACGAAAGCATCAAGCATCGTTTTGCAATGTTTATGCTCGGAGCGGAA 910
Db 853 CCATCACAATATGCAACTAAAGCCATTTTAAATGACATGTTTCTCGGGGAAACAGAA 912
QY 911 ACAACTTCGAAAGCTACAGATGGGTAATGGCAGAGCTGATGAAAAATCAACTGAATCA 970
Db 913 ACTTCATCAACAACAAATTAACGGGCGCATGGTAGAACTGATGAAAAATCAAGTGATTC 972
QY 971 AGAAAAACCAAGAAAGTTAGACAAAGTATTTGGTGAATGGGAAAGTTGATGATCA 1030
```

```
Db 973 GCGAAGCTCAAGCAGAGGTAAAGAGTCTTCAAGGGAAAGAACTTTTCGATGAAGAT 1032
QY 1031 AGATTTTCATGATTTGAAATCTTCAAGTTAGTGTAAAGAAACTCTTAAGATTAATCTCT 1090
Db 1033 GATATCGAGGAGCTGAATTTACCTTTAAGTTAGTCAATTTAGAGAACTTTTAAGATCTCACCT 1092
QY 1091 CCGGTT---GTCCTTGTATTCGAGGAGTGTAGAGAAACCAACACGAAATTTGATGGATGAA 1147
Db 1093 CCACCTTCACCTTTTGTCTTCAAGAGAAATGTCGAGAGAAACAGAAATAAATGGCTACACT 1152
QY 1148 ATTCATCGAAGACTCGAATTTGTGAATGTTGTGAATCTTGGGCGATAGGAAGAGATCCTAATACT 1207
Db 1153 ATTCTTTAAATACCAAGTCAATAGTTAAAGTTTGGGCTATTGGAGAGATCCAAATAT 1212
QY 1208 TSGTCGGAACCTGGAAAGTTTAAACCCAGAAAGGTTTAAAGATTTGCAATTTGATTAATA 1267
Db 1213 TGGGATGATGAGAAAGCTTTTAAGCCTGAGAGATTTGAACATAACTCTTTTGAATTTTGTCT 1272
QY 1268 GGGAGCAGATTTGAACTGGTACCAATTTGGTCAGGAAAAAGAAATATGCTTGGCATTACT 1327
Db 1273 GGCATTAATTTTGAATATCTTCTTTGGTATGGAAGGAGATTGGCCCGGAATATCA 1332
QY 1328 TCAGCTATTACCAATTTGGAGTATGTCATTATAAATCTATTATATCATTTTAAATTTGGGAA 1387
Db 1333 TTTGGTTAGCTAATGCTATCATCTCCATGGCTCAATTTGTTGATCATCTTCGATTGGAGA 1392
QY 1388 CTGCGCGATGGAATTTACACCTCAACACATTTGATATGATGATGAGCTATTGGCGGTGCTCTC 1447
Db 1393 CTCTCTACTGGGTGCGACCAAAATGACTTTGAATTTGACT-AGTTAGCTGGAGTAACACTACT 1451
QY 1448 AGGAAAAAATAGATCTTAAGTTGATTTCTTATTCATAT 1486
Db 1452 GGTAGGAAAAAGAGACCTTTACTTGAATTTTCTACTCTCTTAT 1490

RESULT 6
US-08-948-564-3
; Sequence 3, Application US/08948564
; Patent No. 6121512
; GENERAL INFORMATION:
; APPLICANT: Siminezky, Balazs
; APPLICANT: Dewey, Ralph E.
; APPLICANT: Corbin, Frederick T.
; TITLE OF INVENTION: No. 6121512el Cytochrome P-450 Constructs and
; TITLE OF INVENTION: Methods of Producing Herbicide-Resistant Transgenic Plants
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Virginia C. Bennett
; STREET: PO Box 37428
; CITY: Raleigh
; STATE: No. 6121512th Carolina
; COUNTRY: USA
; ZIP: 27627
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/948,564
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Virginia C.
; REGISTRATION NUMBER: 37,092
; REFERENCE/DOCKET NUMBER: 5051-409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-854-1400
; TELEFAX: 919-854-1401
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1691 base pairs
```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 16..1545
; US-08-948-564-3

Query Match 17.8%; Score 308; DB 3; Length 1691;
Best Local Similarity 53.9%; Pred. No. 4.6e-73;
Matches 769; Conservative 0; Mismatches 620; Indels 39; Gaps 5;

QY 113 CTCCAGGCGCCATGGAAGTTTCCATCATAGGTAAATCTTCCCTCATTTTACTCACTTCT 172
DB 145 CCCCCAGGACCAAGGACACTACCTCTCATAGGAAACATACACACAGATTGTGGC---TCA 201
QY 173 GATCTAGGCCCATGAACGTTTATAGAGCCTTGGCTGCAAAATTTATGACCTGTATAGAGCTTT 232
DB 202 CTGCCGGTTTCACTACTTAAATAATTTGGCAGATAAGTATGTTCCATTTAATGCATCTA 261
QY 233 CAAATTTGGCCAAATTTCACTGTTGTCAATTTCTTCACTGTAAGCAGCCAAAGAGGTTATG 292
DB 262 AAACATAGGAGAGGTGTCACACATCATAGTCACCTTCCCCAGNAATGGCCCCAAGAGATTATG 321
QY 293 AAAAATCAGCGCTGATGCCCTTTCGCCCAACGCCCTTATCGTCTTGGACGCACAGATTGTGTTT 352
DB 322 AAGACACATGATCTCAACTTCTGTATAGCCAGACTTTGTATTGTCTAGATAGTTTCT 381
QY 353 TATAATCGGAAGATGTTGTTTGTGTTTATATGAGATCACTGGAGGCAGATGAAGAAA 412
DB 382 TACAACGGTTCGGCAATTCCTTCACTCAACATGGAGACTATTGGAGGCCAATTAAGAAAG 441
QY 413 ATTTGATATCTGAAATTTCTGAGTGCACAAAAGTTCAATCTCCAGGTTAATCCGAGAG 472
DB 442 ATATGCACAGTAGAGTTACTACAGCAAAAGCGCGTGCAGTCTTTTCGGTCCATAAGAGAA 501
QY 473 GAAGAAATGGAGGATGCCATCATCTCTCCGTTTCA-----AAGCCGATCT 520
DB 502 GAGGAGTGGCAGACTAGTTAAATAATAGTGCNACTGCAAGTGAAGAGGGGGTCTC 561
QY 521 CCGGTCAATATTACAAAGATCAATTTATGCGATTAATTTTCGATCATGATAAGAACATCC 580
DB 562 ATTTTAAATCTCACCAGAGCATTTACTCAATGACTTTTGGGATAGCGGCAGCGGCT 621
QY 581 GTTGTAAATTGAACAAAGAGATTTGCTGAGTGTTCGCGATGCGATCAATGAGCA 640
DB 622 TTGTGTAAGAGCAGATACCAACAAGTGTTCATATCAAAATGATGATAAACAATTTGATG 681
QY 641 GCGACGAGTTTGGCACCCGACACGCTTTTCGACGTGGAAATTTACTTCACTATATCATTT 700
DB 682 CTTCTGGGAGGTTTCTGTTGCTGATCTCTATCTCTCTAGTAGAGTGTTCATATGATG 741
QY 701 GGAGTGTAGTCAAAACCCAGCGGTTTGCATCAGGAGATTGACGATATATCTTGAAGAGATT 760
DB 742 GGGGGCAGCGGGAACCTTGAAAAAGTGCATAGATGACAGATAGGTTGTGCAAGACATC 801
QY 761 CTTAATGACACAAAG-----CCAAATAGCCTTTTGAAGCGGATACATTA 805
DB 802 ATCGACGAGCAAAAATAAGAACAGACGAGGAGCGTGAAGCAGTGGAAATCTTA 861
QY 806 ATGGATGTTTCTATGAATCTTCAAAAAAATGGAACGTTTCCAGTGCAGTGAACAAACGAA 865
DB 862 GTTGATGTTCTCTCAAGTTTCAAAAGGATCGGA-----ATTTCGCTTGTCTGATGAC 915
QY 866 AGCATCAAGACATCCGTTTGGCAATGTTTATGTCGGGAGCGAAACAACTTCGAAAGCT 925
DB 916 AACATTAAGCCGTCATCCAGGACATATTCAATTTGGTGGAGGCGAAACATCATCTTCTGTT 975
QY 926 ACAGATGGTAAATGGCAGAGCTGATGAAAAATCCAACTGAATTAAGAAAGCACAAGAA 985
DB 976 GTGGAATGGGGATGTCAAGATTGATAAGAAAACCCGAGGGTGTGGAAGAAAGCAAGCA 1035

QY 986 GAAGTTAGACAAAGTATTTGGTGAATGGCAAGTTGATGAATCAAGATTTTCATGATTTG 1045
DB 1036 GAGGTGAGAGAGTGTATATAGCAGGATATGTGGATGAGACAAATTTGCCAATTTG 1095
QY 1046 AAATTTCTCAAGTTAGTGGTTAAAGAAACTCTAAGATTACATCTCCGGT---TGTCTTG 1102
DB 1096 ATATACTTTAAAGTCCATCATCAAGAAACCATAGGTTTACATCCACTGTGCGCAATTTG 1155
QY 1103 ATTCCGAGGAGTGTAGAGAAACACACGAATTTGATGGATATGAATTTTCATCCGAACACT 1162
DB 1156 GTTCTTAGAGTAGTAGAGAAAGGTGCCAAATCAATGGATATGAGATACCCCTCTAAGACT 1215
QY 1163 CGAATTTGTTGTAATGCTTTGGCGCATAGGAAGAGATCTCTAAATACATTTGGTCGGAACCTGGA 1222
DB 1216 AGATCATTTATCAATGCTTTGGGCATTTGGAAGGAATCTTAAGTATTTGGGGTGAACCTGAG 1275
QY 1223 AAGTTTAAACCCAGAAAGGTTTAAAGATTGTGCAATTTGATTAATAAGGGACGACATTTGAA 1282
DB 1276 AGTTTAAACCTGAGAGGTTTCTTAATAGCTCCATTTGATTTAGGGGCACAGACTTTGAA 1335
QY 1283 CTGCTACCAATTTGGTGCAGGAAAGAAATATGCTCTGGGCATTTACCTCAGCTATACCAAT 1342
DB 1336 TTTATCCCCATTTGGTGTGGAAGGAGATCTGCCCGGCATTTACATTTGCCATACCAAC 1395
QY 1343 TTGGAGTATGTCATTATAAATCTATTATATCAATTTTAAATTTGGAACTGGCCGATGGAAT 1402
DB 1396 ATTGAGTTGCCACTTGTCTCAGTTACTTTACCACTTTGATTTGNAAGTTTCCCAATAAATG 1455
QY 1403 ACACCTCAACACTTGATATGACTGAAGCTATTGGCGGTGCTCTCAGGAAAAAATAGAT 1462
DB 1456 AAGAATGAAGAACTTGACATGACGAGTCAATGGAATTTACTTTACGAAGACAAATGAC 1515
QY 1463 CTTAAGTTGATCTCTTATTTCCATATCAAGTTAGCTTAGGCTCAATATT 1510
DB 1516 CTCTGCTTGATTTCCCACTACTCGTCTACCTTAAATTTGATGAACAAT 1563

RESULT 7

US-09-292-768-5
; Sequence 5, Application US/09292768
; Patent No. 6194185
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Lupien, Shari L
; APPLICANT: Karp, Frank
; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR THE PRODUCTION OF
; FILE REFERENCE: wbur13463
; CURRENT APPLICATION NUMBER: US/09/292,768
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 08/881,784
; EARLIER FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1515
; TYPE: DNA
; ORGANISM: Mentha piperita
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1494)
US-09-292-768-5

Query Match 14.6%; Score 253.2; DB 3; Length 1515;
Best Local Similarity 50.5%; Pred. No. 2.9e-58;
Matches 671; Conservative 0; Mismatches 648; Indels 9; Gaps 2;
QY 199 CTTGGCTCAAAATTTATGACCTGTTTATGAGTCTTCAAAATTTGCCAAGTTTTCAGCTGTTCT 258
DB 186 CTTGGCGAAGAGTACGGCCCCCGTGGCCCACTGACGCTGGGTGAGGTGTTTCTCCGCTGT 245
QY 259 CATTCTTTCACTGAAGCAGCAAGAGGTTTATGAAAACTCAGGCTGATGCTTCGCCCA 318

Db 246 CTTTTCGTGCGGGAGGCGAAGGAGGCGGATGAGTGGTAGACCGGCGGTGCGGAA 305
Qy 319 ACGCCCTATCGTCTTGGAGCGACAGATTGTTTATTAATCGGAAGAGTCTCTGTTTGC 378
Db 306 CCGGTTCCAGAGCATCGGACGAGGATCATGTGGTAGCAACAGGAGCATCATCTTCAG 365
Qy 379 TTCTATATGAGATCACTGAGGCGAGATGAAGAAATTTGGATACATTGTAATTTCTGAGTGC 438
Db 366 CCCCTACAGCGAGCATCGGCGCAGATGCGCAAGATCTCGGTCTCCGAGCTCTCTCCTC 425
Qy 439 CAAAAAGTTCAATCCTCCAGTTAATCGAGAGGAAGAAATGGAGGATGCCATCACCATT 498
Db 426 CCGCAACGTCCGCTCTCTCGGCTTCATCCGCGAGGACGAGGTGTCGCGCTCTCCGCCA 485
Qy 499 CCTCGGTTCCGAAGCGGATCTCCGCTCAATATTAACAAAGATCATTTATGCAATTATAAT 558
Db 486 CCTCGCTCTCGGAGGCGGCGGTGACATGACGAGAGGATAGAGACGCTGACGTG 545
Qy 559 TTCGATCATGATGAAGAACATCCGTTGGTAAATGTTAA---GCAAAAGAAAGATTCCTGAG 615
Db 546 CTCCATCATCTCAGGGCGGCTTCGGGAGCGTATCAGGGACAAACGCGGAGCTGTTGGG 605
Qy 616 TGTTCCGATGACGATCAATGAGCGAGCGAGGTTTTCGACCGCAGACGCTTTTCCGAC 675
Db 606 GCTGTCAGGACGCGCTCAGCATGGCTCGGGGTTGAGCTCGCGGACATGTTCCCTC 665
Qy 676 GTGAAATTTACTTCACTATATCATTTGGAGCTGAGTCAAAACCCAGCGGTTGTCATCAGGA 735
Db 666 CTCGAGCTCTCAACCTCTCTGTGNAACAGAGCAAGCTCTGGAGGATGCGCGCG 725
Qy 736 GATTCAGATATACCTGTAAGAGATCTTAAATGAAACAAAGCCAATAAG-----CCTTT 789
Db 726 CGTCGACACCATCTCTCGAGGCCATCGTCGAGCAGCAAGTTTCAAGAAAGCGCGGAGTT 785
Qy 790 TGAAGCGGATCTTAATGAGTGTCTATTGATCTTCAAAAAATGNAAGTCCAGT 849
Db 786 CGGCGGCGAGGACATCATCGAGCTCTCTTCAAGATGAGAGGAGCCAGCATCAAGT 845
Qy 850 GCCAGTGACAAAGCAAGCATCAAAAGCATCCGTTTGGCAATGTTTACTGCGGGAGCGGA 909
Db 846 CCCCATCACCACCACTCCATCAAGCCTTCATCTTCGATAGTTCTCAGCAGGAGTGA 905
Qy 910 AACATCTTGAAGCTACAGATGGGTAAATGCGAGAGCTGATGAAAAATCCAACTGAAT 969
Db 906 GACATCTCTCAACACACCATCTGAGTGTGGTGGGAGCTGATGAGGAAACCCGCGAGTGA 965
Qy 970 AAGAAAGCACAAAGAAAGTTAGACAAGTATTTGGTGAATGGGAAAGTTGATGAATC 1029
Db 966 GCGGAAAGCGCAGGCGGAGGTGAGCGGCATCTGAAGGAGAGACGAACTGGGAGTGA 1025
Qy 1030 AAGATTTTCATGATTTGAAATCTTCAAGTTAGTGTAAAGAACTCTTAAGATTAATCC 1089
Db 1026 TGATGTGCAAGAGCTTAAGTACATGAATCGGTGGTGAAGGAGACGATGAGGATGACCC 1085
Qy 1090 TCCGTTGTCTGATTCGAGGAGGTGAGAGAAACACACAAATGATGGATATGAAT 1149
Db 1086 TCCGATCCGTTGATCCGAGATCATGCAAGAAAGTTCGGTGTAAACGGGTATACGAT 1145
Qy 1150 TCATCCGACATCGAATTTGTTGTAATGCTTTGGCGATGAGGAGATCCTAAATCTTG 1209
Db 1146 TCCGAACAGGCGAGATCATGATCAACGTCGTGTCATGGCAGGAATCCTCTACTG 1205
Qy 1210 GTCCGAACCTGGAAAGTTTAAACCGAAGAGTTTAAAGATTTGCAATTTGATAAAGG 1269
Db 1206 GGAAAAACCGATACCTTTTGGCCGAAAGTTTGAACCAAGTTTCAAGGATTTTCATGG 1265
Qy 1270 GAGGACATTTGAATGATGATACATTTGGTCAGGAAAGAAATATGTCCTGGCATTAATC 1329
Db 1266 AAATGATTTTCGATGTCCTCGGTCGAGCGGAGAAAGATCTGCCCGCGGTTGAATTT 1325
Qy 1330 AGCTATTACCAATTTGGAGTGTGTCATTATAATCTATTATATCTTTTAAATTTGGAACT 1389
Db 1326 CCGGTCGGCAAGGTTGAGGTTCCATTTGCGGCGAGCTTCTTTACCAGTTCGATGGAATTT 1385

Qy 1390 GGCGCATGGAATTACACCTCAAAACACCTGATATGACTGAAGCTATTGGCGGTGCTCTCAG 1449
Db 1386 GGCGGAAGGAATGAACCTTCTGATATGACATGCTGAGCGGAAGCCTTACCGGAAT 1445
Qy 1450 GAAAAAATAGATCTTAAGTTGATTCCTTATTTCCATATCAAGTTAGCTTAGCTCAAAATAT 1509
Db 1446 ACTAAGAACAAATCTTCTTCTTGTTCACACCCCTACGATCCTCTCATGATCAATCAT 1505
Qy 1510 TTCTTGAT 1517
Db 1506 CTTTGTCT 1513

RESULT 8

US-08-881-784-5
; Sequence 5, Application US/08881784
; Patent No. 6083731
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B.
; APPLICANT: Lupien, Shari L.
; APPLICANT: Karp, Frank
; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR
; TITLE OF INVENTION: THE PRODUCTION OF LIMONENE HYDROXYLASES
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen, O'Connor, Johnson and Kindness
; ADDRESS: PLLC
; STREET: 1420 Fifth Avenue, Suite 2800
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/881.784
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Shelton, Dennis K.
; REGISTRATION NUMBER: 26,997
; REFERENCE/DOCKET NUMBER: WSUR19777
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 224-0718
; TELEFAX: (206) 224-0779
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1762 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Mentha spicata
; INDIVIDUAL ISOLATE: cDNA encoding
; INDIVIDUAL ISOLATE:
; IMMEDIATE SOURCE:
; CLONE: pSM12.2
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 558..1212
; OTHER INFORMATION: /product= "Probe LH-1 (Figure 4A)"
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 39..538
; OTHER INFORMATION: /product= "Probe LH-2 (Figure 4A)"
US-08-881-784-5

Query Match 14.2%; Score 245.4; DB 3; Length 1762;

		Best Local Similarity 50.5%; Pred. No. 4e-56;		Matches 706; Conservative 0; Mismatches 676; Indels 15; Gaps 4;	
QY	102	AACAGAACTCCACCTCCAGGCGCATGGAAGTTTCCTATCATAGTAACTCTTCCTCATTTAT	161		
DB	111	AACAAAACCTACTCTCGAGCCCTCGGAAGCTCGCGGTGATCGGCCACCTCC-----ACT	164		
QY	162	TACTCACTTCTGATCTAGGCCATGAACGTTTTAGAGCCTTGGCTCAAAATTTATGACCTG	221		
DB	165	TCCTGTGGGAGGGCTTCCAGCAGCAGCTGTTTAGGAGCATGCCAGAGTACGGGCCG	224		
QY	222	TTATGAGTCTCAAAATGGCCAAAGTTTCAGCTGTTGTGTCATTTCTCAGCTGAAGAGCCA	281		
DB	225	TGGCCACATGTCAGCTGGGAGAGTGTACTCGGTGGTGTGCTCGTGGCGGAGGAGCGA	284		
QY	282	AAGAGTTATGAAAACTCAGGCTGTATGCTTCGCCCAACGCCCTATCGTCTTGGAGCGAC	341		
DB	285	ACGAGCGATGAAGTGTGAGCCCGAACTTCGCCGACCGGTTCAGCGGCATCGGGTCCA	344		
QY	342	AGATTGTGTTTTATAATCGGAAGATGTCTTGTGTTTTCATATGGAGATCACTGGAGGC	401		
DB	345	GGACCATGTGTGACGACAAAGATGACATCATCTTCAGCCCTTACAAAGATCACTGGGCC	404		
QY	402	AGATGAAGAAAATTTGGATACCTTGAATTTCTGAGTGCCAAAAAAGTTCAATCTCCAGGT	461		
DB	405	AGATCGGAGGATCTCGGTGACAGAGCTGCTGAGCCCGAAGAACGTCAGGTCTCTCGGT	464		
QY	462	TAATCCGAGAGGAAGATGGAGGATGCCATCACATTCCTCGTTCGAAAGCGGATCTC	521		
DB	465	ACATAAGGAGGAGGATCGAGCGCCTCATCGGCTGCTCGGGTGTGCGGGGAGCGC	524		
QY	522	CGGTCAATATTACAAAGATCATTTATGGCATTATAATTTTCGATCATGATAAGAACATCG	581		
DB	525	CGGTGACGTGACGAGGAGGTGTGAGATGTGCTGTGCTGTGTCAGGGGGCGGT	584		
QY	582	TTGTGTAATTTGAAGCAAAAGAAAGATGCTG--AGTGTGCGGATGAGTCAAT--GAGG	638		
DB	585	TCGGGAGTGTCTCAAGGACCAAGGTTCTGTGGCGAGTTGGTGAAGGATCGCTGGCAT	644		
QY	639	CAGCGACGATTTTGGCCAGCAGACGCTTTTCGACGTGGAAATTACTTCACTATATCA	698		
DB	645	TGGCGTCCGGGTTTGAAGTGGGATCTCTACCTTCTCTCATGGCTCTCTCAACCTGCTTA	704		
QY	699	TTGAGCTGAGTCAAAACCCAGCGCTTTGTCATCAGGAGATTGACGATATACITGAAGAGA	758		
DB	705	GCTTGAACAGTACAGTTGACAGAGATGGCGCGCGCTCGATCACATCTTGTATGGT	764		
QY	759	TTCTTAATGAACACAAAGCAAATAAG-----CTTTTGAAGCGGATTAATTTAATGGATG	812		
DB	765	TCCTGGAGGAGCATAGGGAGAAAGAGCGCGGAGTTTGGAGCGGAGACATCGTCGACG	824		
QY	813	TTCTATTGAATCTTCAAAAAATGNAAGTTCCAGTGCCAGTGCCAGTGACAAACGAAAGCATCA	872		
DB	825	TTCTTTTCAGGATGACAGAGGCGACGACATCAAAATTTCCATTACTTCCAAATTCGATCA	884		
QY	873	AAGCATCCGCTTTTGAATGTTTACTGCGGGAGGCAAACTTCGAAAGCTACAGAT	932		
DB	885	AGGGTTTCATTTTCACACCTTCTCGCGGAGCTGAAACGTTCTGACGACCATCTCAT	944		
QY	933	GGGTAAATGGCAGCTGTATGAAAAATCCAACTGAACCTAAGAAAAAGCAAGAAAGTTTA	992		
DB	945	GGCGGTGTGCGAACTGATGAGGAATCCGGCGAAGATGGCCAAAGTGGCGGCGGAGGTA	1004		
QY	993	GACAGATTTTGGTGAATGGGAAAGTTGATGAATCAAGATTTTCATGATTTGAATTTCT	1052		
DB	1005	GAGAGCGCTCAAGGAAAGACAGTCTGTGATTTGAGCGAGGTGCAAGAGCTAAATAATACC	1064		
QY	1053	TCAAGTTAGTGTGTTAAAGAACTCTAAGATTACATCTCCCGTGTGCTTTGATTCCGAGG	1112		
DB	1065	TCAGATCGGTGTAAAGGAGATCTTGAGGCTGCACCTCTTTCATTAATCCCAAGAC	1124		
QY	1113	AGTGTAGAGAAACAAACAGAAATTTGATGGAATTAATTCATCCGAACACTCGAATTTGTG	1172		

DB	1125	AATCAGGGAAGAATCCGAGTTAACGGGTACACGATTCGGGCCAAAACTAGAAATCTTCA	1184		
QY	1173	TGAATGCTTTGGCGATAGGAAGAGATCCTAACTACTTGTTCGNAACCTGGAAGTTTAAAC	1232		
DB	1185	TCAGAGTCTGGGCTATCGGAAGGATCCCAATACTCTGGGAAGATCCCGACACCTTCGGCC	1244		
QY	1233	CAGAAAGGTTTAAAGATTGTGCAATTGTGCAATTGATTATAAAGGAGCAGACATTTGAACTGGTACCAT	1292		
DB	1245	CTGAGAGATTTCGATGAGTTTCCAGGGAATTTTCATGGGAACGATTTTCGATTCATCCCAT	1304		
QY	1293	TTGTGTCAGGAAAAAGATATGTCCTCGCATTACTTTCAGCTATTACCAATTTGGAGATG	1352		
DB	1305	TCGGGGCGGTGAAAGAACTCTGCCCGGTTTACATTTCCGGCTGGCAAAATGTTGAGATCC	1364		
QY	1353	TCATTATAAATCTATTATATCATTTTAAATTTGGAACTGSCCGATGCAATTCACCTCAAA	1412		
DB	1365	CATTGGCGAATGCTCTACCACTTCGACTTGGAAATTCGCAAGGATGACTGATGCGG	1424		
QY	1413	CACTTGATATGACTGAAGCTATTGGCGGTCTCTCAGGAAAAAATAGATCTTAACTTGA	1472		
DB	1425	ACTTGGACATGACGGAGACCCCGAGTCTTCTGGGCCAAAAAGAAATGTTTGTCTTG	1484		
QY	1473	TTCTATTCCATATCAA 1489			
DB	1485	TTCCACACCTCTATAAA 1501			

RESULT 9

US-09-292-768-1

; Sequence 1, Application US/09292768

; Patent No. 6194185

; GENERAL INFORMATION:

; APPLICANT: Croteau, Rodney B

; APPLICANT: Lupien, Shari L

; APPLICANT: Karp, Frank

; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR THE PRODUCTION OF LIMONENE HYDROXYLASES

; FILE REFERENCE: wsl13463

; CURRENT APPLICATION NUMBER: US/09/292,768

; CURRENT FILING DATE: 1999-04-14

; EARLIER APPLICATION NUMBER: 08/881,784

; EARLIER FILING DATE: 1997-06-24

; NUMBER OF SEQ ID NOS: 70

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 1762

; TYPE: DNA

; ORGANISM: Mentha spicata

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (20) .. (1507)

US-09-292-768-1

Query Match 14.2%; Score 245.4; DB 3; Length 1762;

Best Local Similarity 50.5%; Pred. No. 4e-56;

Matches 706; Conservative 0; Mismatches 676; Indels 15; Gaps 4;

QY	102	AACAGAACTCCACCTCCAGGCGCATGGAAGTTTCCTATCATAGTAACTCTTCCTCATTTAT	161		
DB	111	AACAAAACCTACTCTCGAGCCCTCGGAAGCTCGCGGTGATCGGCCACCTCC-----ACT	164		
QY	162	TACTCACTTCTGATCTAGGCCATGAACGTTTTAGAGCCTTGGCTCAAAATTTATGACCTG	221		
DB	165	TCCTGTGGGAGGGCTTCCAGCAGCAGCTGTTTAGGAGCATGCCAGAGTACGGGCCG	224		
QY	222	TTATGAGTCTCAAAATGGCCAAAGTTTCAGCTGTTGTGTCATTTCTCAGCTGAAGAGCCA	281		
DB	225	TGGCCACATGTCAGCTGGGAGAGTGTACTCGGTGGTGTGCTCGTGGCGGAGGAGCGA	284		
QY	282	AAGAGTTATGAAAACTCAGGCTGTATGCTTCGCCCAACGCCCTATCGTCTTGGAGCGAC	341		
DB	285	ACGAGCGATGAAGTGTGAGCCCGAACTTCGCCGACCGGTTCAGCGGCATCGGGTCCA	344		


```
Db 465 ACATAAGCAGGAGGAGATCGAGCGCCTCATCGCGTGTCTCGGGTGTCTCGGGGGAGCGC 524
QY |||||
Db 522 CGGTCAATATTACAAAGCATATTATGGCATTATAATTCGATCATGAATGAACATCCG 581
Db 525 CGGTGACGTGACGAGAGAGGTGCGAAGATGTGCTGTGCTGTGTCAGGCGCGCGT 584
QY 582 TTGGTAATTGAAGCAAAAGAAAGATTGCTG--AGTGTTCGCGATCGAGTCAAT--GAGG 638
Db 585 TCGGGAGTGTGCTCAAGACACAGGTTTCGTTGGCGAGTTGGTGAAGGAGTCTCGTGCAT 644
QY 639 CAGGACGAGTTTGGCACCGCAGACGCTTTTCGAGCTGGGAAATTAATTCATATATCA 698
Db 645 TGGCGTCCGGTTTGAAGTGGCGGATCTCTACCTTCTCATGGCTCTCAACCTGCTTA 704
QY 699 TTGGAGCTGAGTCAAAACCCAGCGTTTTCATCAGGAGATTGACGATATATCTTGAAGAGA 758
Db 705 GCTTGAACAGTACAGGTTTCAGAGGATGCGCGCGCTCGATCATCTTTGATGGGT 764
QY 759 TTCTTAATGAACACAAAGCCAAATAAG-----CCTTTGAAGCGGATAACTTAATGGATG 812
Db 765 TCCTGGAGGAGCATAGGAGAGAGAGCGCGAGTTTGGAGCGGAGGACATCGTCGAGC 824
QY 813 TTCTATTGAATCTTCAAAAATAATGGAACCTTCAGTGCCAGTGAACAAAGCAAGCATCA 872
Db 825 TTCTTTTCAGGATGTCAGAGGCGCAGACATCAAAATTTCCATTACTTCCAAATTGCATCA 884
QY 873 AAGCATCCGTTTTCGAAATGTTTACTGCGGGAGCGAACAACCTTCGAAAGCTACAGAAT 932
Db 885 AGGTTTCATTTTCGACACCTTCTCCGCGGAGCTGAAAGCTTTCGAGACCATCTCAT 944
QY 933 GGTAATGCGAGAGCTGATGAAAAATCCAACTGAACCTAAGAAAAGCACAAAGAAAGTTA 992
Db 945 GGGCGTTGTCGAACTGATGAGGAATCCGCGAAGATGGCAAGTGGCAGCGGAGTAA 1004
QY 993 GACAGTATTGGTGAATGGGAAAAGTTGATGAATCAAGATTTTCATGATTTGAATCT 1052
Db 1005 GAGAGCGCTCAAGGGAAGACAGTGTGATTTGAGCGAGGTGCAAGAGCTAAAAATACC 1064
QY 1053 TCAAGTTAGTGTAAAGAACTCTAAGATTACATCTCCGGTGTCTTGAATCCAGGG 1112
Db 1065 TGAGATCGGTGTAAAGGAGACTCTGAGGCTGACCCCTCCCTTTCATTAATCCCAAGAC 1124
QY 1113 AGTGTAGAGAAAACACACGAAATTTGATGGATATGAAATTCATCCGAACACTCGAATTTG 1172
Db 1125 AATCCAGGAGAGATCGGAGTTAACGGGTACACGATTCCGGCAAACTAGATCTTCA 1184
QY 1173 TGAATCCTTTGGCGATAGGAAGAGATCTTAATCTTGGTGGAACTGGAAAGTTTAAACC 1232
Db 1185 TCAACGCTCTGGCTATCGGAAGGATCCCAATACTTGGGAAGATCCGACACCTTCGCGC 1244
QY 1233 CAGAAAGTTTAAAGATTGCAATTGATTAATAAGGAGCAGACATTTGAACTGGTACCAT 1292
Db 1245 CTGAGAGATTTCGATGAGTTTCCAGGATTTTCATGGGAACGATTTTCGAGTTTCATCCCA 1304
QY 1293 TTGGTGCAGGAAAAGAAATATGCTTCGCAATTAATTCAGTATTAACCAATTTGGAGTATG 1352
Db 1305 TCGGGCGGTCGAGAACTGCCCCGGTTTACATTTCCGGTGGCAATTTGATATCC 1364
QY 1353 TCATATAAATCTATTATATCAATTTTAAATGGGAACCTGCCCGATGGAATTAACCTCAAA 1412
Db 1365 CATTTGGCGCAACTGCTCTACCACTTCGATGGAAATTTGCCAAGGAATGACTGATGCG 1424
QY 1413 CACTTGATATGACTGAGCTATTGGCGGTGCTCTCAGGAAAAAATAAGATCTTAAGTTGA 1472
Db 1425 ACTTGAATGATGAGGAGACCCAGAGTCTTTCTGGGCCAAAAAAGAAATGTTTCTGTGG 1484
QY 1473 TTCTTATTCATATCAAA 1489
Db 1485 TTCCACACATCTATATAA 1501
```

RESULT 11

US-09-292-768-67

```
; Sequence 67, Application US/09292768
; Patent No. 6194185
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Lupien, Shari L
; APPLICANT: Kaip, Frank
; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR THE PRODUCTION OF
; FILE REFERENCE: waur13463
; CURRENT APPLICATION NUMBER: US/09/292,768
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 08/881,784
; EARLIER FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 67
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: computer-generated nucleic acid sequence encoding
; OTHER INFORMATION: limonene-3-hydroxylase
; NAME/KEY: misc feature
; LOCATION: (1)..(1665)
; OTHER INFORMATION: computer-generated nucleic acid sequence encoding
; OTHER INFORMATION: limonene-3-hydroxylase variant
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19)..(1518)
; US-09-292-768-67
```

```
Query Match 14.1%; Score 243.8; DB 3; Length 1665;
Best Local Similarity 50.2%; Pred. No. 1.1e-55;
Matches 692; Conservative 0; Mismatches 672; Indels 15; Gaps 3;
```

```
QY 199 CTTGCTCAAAATTTATGGACCTGTTATGAGTCTTCAAATTTGCCAAGTTTTCAGCTTTGT 258
Db 204 CTTGCGAAGCAGTACGCGCCAGTGGCGACGCTGCGAGGTGTTCTCCGCTGT 263
QY 259 CATTTCTTCAGCTGAAGCAGCAAGAGGTTATGAAACTCAGGCTGATGCCATCACATT 318
Db 264 GCTCTGTCCTCCGAGGCCACGAGGAGCGATGAAGCTGGTGGACCCGCTTCGCGGA 323
QY 319 AGCCCTATPCTGTTTGGACGACAGATTGTGTTTTATATCGAAAGATGTCTTGTTCG 378
Db 324 CCGGTTTCGAGAGCATCGGAGCGAAGATCATGTGTGACGACACGACATCATCTTCAG 383
QY 379 TTCATATGGAGATCACTGGAGCAGATGAAGAAATTTGGATATCTTAATTTCTGATGC 438
Db 384 CCGCTACAGCTGCACTGGCGCCAGATCGGGAAGATCTGCGTCTCCGAGCTCTCAGCGC 443
QY 439 CAAAAAAGTTCAATCTCCAGGTTAATCCGAGAGCAAGAAATGGAGGATGCCATCACATT 498
Db 444 CCGAAGCTCCGCTCTTTCGGCTTCATCAGCAGACGAGGTGTCCTCCGCTCTCTCGCCA 503
QY 499 CTTCCGTTTCAAGGCC-----GGATCTCCGCTCAATATTACAAAGATCATTTATGGCAT 552
Db 504 CTTCCGCTCTCTCCGCGCGCGGGGAGCGCTGACCTCAGGAGCGGATAGCAGCGCT 563
QY 553 TATAATTTGATCATGATAAGAACATCCGTTGGTAAATTGTAA---GCAAAAAGAAAGATT 609
Db 564 GAGTGCTCCATCATCTCGAGGGCGCGTTCGGGAGCGGTGATCAGGAGCCACGAGAGCT 623
QY 610 GCTGAGTGTTCGCGATGCAATGAGGACGACGAGTTTGGCACGACGAGCTTT 669
Db 624 GGTGAGCTGTGAAGGAGCGCTCAGCATGCGGTTCGGGTTTCGAGCTCCGCAATGTT 683
QY 670 TCCGAGCTGGAATTTACTTCACTATATCATTTGAGAGCTGAGTCAAAACCCAGCGTTTGA 729
Db 684 CCCCCTCCCAAGCTCCTCACTTGTCTGCTGGAAACAAGCAAGCTGTGGAGATGCG 743
```

QY 730 TCAGGAGATTGAGTATACCTTGAAGAGATCTTAAATGAACACAAGCAATAAG----- 784
Db 744 CGCGCGTGCAGCGCCATCTCGAGGCCATCGTGAGGAGCAACAGCTCAAGAAGAGCGG 803
QY 785 -CCTTTTGAAGCGGTAACCTTAATGATGTTCTATTGAATCTTCAAAAAATGGAACGT 843
Db 804 CGAGTTTGGCGGAGGACATATTGACGTACTCTTTAGGATGCAAGAGGATAGCCAGAT 863
QY 844 TCCAGTGCAGTGACAAACGAAAGCATCAAGCATCCGTTTTTGCAAATGTTTACTGCCGG 903
Db 864 CAAAGTCCCATCACCAACCAAGCCATCAAGCCCTTCTTCGACACGTTCTCAGCGG 923
QY 904 GAGCAACAACTTCCAAAGCTACAGATGGTAAATGSCAGAGCTGATGAAAAATCCAC 963
Db 924 GACCGAGACATCATCAACACCAACCCCTGTGGGTGATGGCGGAGCTGATGAGGAATCCAGA 983
QY 964 TCAACTTAAGAAAACACAGAAGAAGTTAGACAAGTATTGTTGTAATGGGAAAAGTTGA 1023
Db 984 GGTGATGGCAAAAGCGAGCGGAGGTGAGCGCGCTGAAGGGGAAGACGGACTGGGA 1043
QY 1024 TGAATCAAGATTTCAATTTGAAATTTCTTCAAGTTAGTGGTTAAAGAAACTCTAAGATT 1083
Db 1044 CGTGACGACGTCGAGGAGCTTAAAGTACATGAAATCGGTGGTGAAGGAGACGATGAGAT 1103
QY 1084 ACATCTCCGTTGTTCTGATTCGAGGAGGTGTAGAGAAACACACGNAATTGATGGATA 1143
Db 1104 GCACCCCTCGATCCCGTTGATCCCGAGATCATGSCAGAGAAGAAATCGGAGGTCAACGGTA 1163
QY 1144 TGAATTCATCCGAACACATCGAATTTGTTGTAATGCTTGGGCGATAGGAAGAGATCTCTAA 1203
Db 1164 CAGGATTCGAATAGCCAGAAATCATGATCAAGTGTGTCCTGAGGATGAGGATCTCTCT 1223
QY 1204 TACTTGGTCGAACTCGGAAAGTTTAAACCCAGAAAGTTTAAAGATTTGCAATTTGATTA 1263
Db 1224 CTACTGGGAAAACCCGAGACCTTTTGGCCGAAAGTTTGAACCAAGTCTCGAGGATTT 1283
QY 1264 TAAAGGAGGACATTTGAAGTGGTACCATTTGGTCGAGGAAAAGAAATATGTCCTGGCAT 1323
Db 1284 CATGGAAACGATTTTCGATTCATCCCATTTTGGAGCTGGAAGAAGAAATCTGCCCCGGTT 1343
QY 1324 TACTTCAGCTATTACCAATTTTCGAGTATGTCATTAATATATATATATATATATTTAATTG 1383
Db 1344 GAATTTCCGTTGGCAATGTTGAGTCCCATTTGGCAGCATCTTTTACCACCTTCGACTG 1403
QY 1384 GGAAGTGGCCGATGGAATTTACCTCAACACTTGAATGATGATGATGATGATGATGATGATG 1443
Db 1404 GAAATTGGCGGAAGGAATGAACCTTCCGATATGACATGCTGAGGAGCAAGAGCCCTTAC 1463
QY 1444 TCTCAGGAAAATAAGATCTTAAAGTTGATTCCTATTCCATATCAAGTTAGTTAGGCTC 1503
Db 1464 CGGAATAAGAAAAGAAATCTTCTACTCGTTCACACCCCTACGATCTCTCATGATC 1523
QY 1504 AAATATTTCTGATACATAGGAGGTTGAAATATATATATAATAATAATAATAATAATAAGAT 1562
Db 1524 AATTAATCTCTTTAATTTGCTCTTGAATAAAGAGTGATATATATATATATATATATATAT 1582

RESULT 12
US-09-292-768-63
; Sequence 63, Application US/09292768
; Patent No. 6194185
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Lupien, Shari L
; APPLICANT: Karp, Frank
; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR THE PRODUCTION OF
; FILE REFERENCE: wsur13463
; CURRENT APPLICATION NUMBER: US/09/292,768
; CURRENT FILING DATE: 1999-04-14
; EARLIER FILING DATE: 08/881,784
; NUMBER OF SEQ ID NOS: 70

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 63
; LENGTH: 1762
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: computer-generated nucleic acid sequence
; NAME/KEY: CDS
; LOCATION: (20)..(1507)
; OTHER INFORMATION: Computer-generated nucleic acid sequence encoding
; OTHER INFORMATION: limonene-6-hydroxylase variant
US-09-292-768-63

Query Match 14.1%; Score 243.8; DB 3; Length 1762;
Best Local Similarity 50.5%; Pred. No. 1.1e-55;
Matches 705; Conservative 0; Mismatches 677; Indels 15; Gaps 4;
QY 102 AACAGATCCACCTCCAGGCGCATGGAAGTTTCTTATCATAGGTAATCTTCTCATTTAT 161
Db 111 AACAAAACCTACCTCCGAGCCCTCCGAAGCTGCCGGTATCGGCCACCTCC-----ACT 164
QY 162 TACTCACCTTCTGATCTAGGCCATGAACGTTTTTAGAGCCTTGCTCAAAATTTATGACCTG 221
Db 165 TCCTGTGGGAGGGCTTCCCGACGACGTTTGGAGCATAGCCAGAGTACGGCGCGG 224
QY 222 TTATGAGTCTTCAAATTTGGCCAAAGTTTCAGCTGTTGTCAATTTCTTCACTGAAGAGCA 281
Db 225 TGGCGCAGCTGCAGCTGGGAGAAGTGTACTCGGTGGTCTGTCTGCGCGGAGGAGCGA 284
QY 282 AAGAGGTTATGAAAACCTCAGGCTGATGCTTCGCCCAAGCCCTATCTGCTTGGAGGCAC 341
Db 285 AGCAGGCGATGAAGTGTCTGGACCCGAACCTTCGCCGACCGTTCCAGCGCATCGGGTCCA 344
QY 342 AGATTGTTGTTTTAAATCGGAAAGATGCTTTGTTGTTTTCATATATGGAGATCACTCGAGGC 401
Db 345 GGACCATGTGTGACGAAAGATGACATCATCTTCAGCCCTTACACGATCACTGGCGCC 404
QY 402 AGATGAAGAAAATTTGGATACCTTGAATTTCTGAGTGCACAAAAGTTTCAATCTCTCAGGT 461
Db 405 AGATGCGSAGATCTGCGTGACAGAGTGTGAGCCCGAAGAACGTCAGGTCTCTCGGGT 464
QY 462 TAATCCGAGAGAGAATAATGAGAGTGCATCACAATCTCCTCGTTCGAAAGCCGATCTC 521
Db 465 ACATAAGCAGCAGGAGAGATCGAGCGCTCATCTCCGGTCTCGGGTCTGCGGGGAGGCGC 524
QY 522 CGGTCAATATTACAAAGATCATTTATGGCATTAATAATTCGATCATGATAAGAACATCCG 581
Db 525 CGGTGACGTCACGAGGAGGTGTGCAAGATGTCTGTCTGTCTGTCGAGGCGCGGT 584
QY 582 TTGTAAATTTGAACAAAAGAAAGATTTGCTG--AGTGTTCGCGATGAGTCAAT--GAGG 638
Db 585 TCGGAGTGTCTCAAGGACCGAGGTTTCTGTGGCGAGTTTGGTGAAGAGTCTGCTGGCAT 644
QY 639 CAGCAGGAGTTTGGCACCCGACAGCTTTTCCGACGTGGAAATTAATCTCACTATATCA 698
Db 645 TGGCTCCGGTTTGAAGTGGCGGATCTCTACCTTCTCATGGCTCTCAACCTGCTCTA 704
QY 699 TTGGAGCTGAGTCAAAACCCAGGCGTTTGGCATCAGGAGATTGACCATATATCTTGAAGAGA 758
Db 705 CTTGAACAGTACAGGTTGAGAGGATGCGCCCGGCTCGATCAGATCTTGTATGGT 764
QY 759 TTCTTAATGAACAAAGCCCAATTAAG-----CCTTTTGAAGCGGATTAATTTAATGATG 812
Db 765 TCCTGGAGGAGCATAGGGAAGAAAGAGCGCGAGTTTGGAGCGGAGGACATCTGTCGAGC 824
QY 813 TTCTATTGAATCTTCAAAAAAATGGAACGTTCCAGTCCAGTGAACAAAGCAATCA 872
Db 825 TTCTTTTTCAGGATGAGAAAGGCGAGCATCAAAATTTCCCATTAATTTCCCAATTCATCA 884
QY 873 AAGCATCCGTTTTCAAAATGTTTACTCGCGGAGGAGCAAACTTCGAAAGCTACAGAT 932

Db 885 AGGTTTCATTTTCGACACCTTCTCCGCGGAGCTGAAACGCTTCGCGACCATCTCAT 944
QY 933 GGTATATGCGAGCTGATGAAAAATCCAACTGAACCTAAGAAAAGCACAGAGAGTTA 992
Db 945 GGGCTTGTGCGACTGATGAGGATCCGCGAGATGGCCAGGTGCGAGCGAGGTAA 1004
QY 993 GACAAGTATTTGGTGAATGGGAAAGTTGATGAATCAAGATTTTCATGATTTTGAATTTCT 1052
Db 1005 GAGAGCGCTCAAGGGAAGACAGTCTGTGATTTGAGCGAGGTGCAAGAGCTAAAATACC 1064
QY 1053 TCAAGTATGTTTAAAGAACTTAAGATTACATCTCTGGTGTGCTTGTGATTCGAGGG 1112
Db 1065 TGAGATCGGTGTTAAAGGAGACTCTGAGCTGACCCCTCTCTTCCATTAATCCCAAGAC 1124
QY 1113 AGTGTAGAGAAACACACGAATTTGATGATATGAAATTCATCCGAACACTCGAATTTGTTG 1172
Db 1125 ATCCAGGAGAGATTCGAGGTTAACGGGTACAGGATTCGGGCAAACTAGAAATCTCA 1184
QY 1173 TGAATGCTTGGCGGATPAGAAAGAGATCTTAATATCTTGGTTCGGAACCTGGAAGTTTAAAC 1232
Db 1185 TCAAGCTCTGGCTATCGGAAGGATCCCAATACTTGGGAAGATCCCGACACCTTCGCGC 1244
QY 1233 CAGAAGGTTTAAAGATTTGCAATTTGATTAATTAAGGAGACATTTGAACCTGGTACCAT 1292
Db 1245 CTGAGAGATTCGATGAGGTTCGAGGATTTTCATGGGAAACGATTTTCGAGTTTCATCCCAT 1304
QY 1293 TTGGTGCAGAAAAGATATGTCCTGGCATTTACGCTATTACCAATTTGGAGTATG 1352
Db 1305 TCGGGCGGTGCAAGAAATCTGCCCGGTTTACATTTTCGGGTGCGAAATGTTGAGATCC 1364
QY 1353 TCATTTATAATCTATTATATCAATTTTAAATGGGAACCTGGCCGATGGAATTTACACCTCAA 1412
Db 1365 CATTCGGCGCACTGTCTTACCCTTCGACTGGAATTTGCCAAGGAATGACTGATGCGG 1424
QY 1413 CACTTGATATGACTAGACTATTTGGGGTGTCTCAGGAAAAAATAGATCTTAAGTTGA 1472
Db 1425 ACTTGACATGACGAGACCCAGGTCTTTCTGGGCCAAAAAAGAAAAATGTTGCTGG 1484
QY 1473 TTCTATTCCATATCAA 1489
Db 1485 TTCCACACTCTATAAA 1501

RESULT 13

US-08-881-784-8
; Sequence 8, Application US/08881784
; Patent No. 6083731
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B.
; APPLICANT: Lupien, Shari L.
; APPLICANT: Karp, Frank
; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR
; THE PRODUCTION OF LIMONENE HYDROXYLASES
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen, O'Connor, Johnson and Kindness
; ADDRESSEE: PLLC
; STREET: 1420 Fifth Avenue, Suite 2800
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/881,784
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Shelton, Dennis K.

; REGISTRATION NUMBER: 26,997
; REFERENCE/DOCKET NUMBER: WSUR19777
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 224-0718
; TELEFAX: (206) 224-0779
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1665 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Mentha x piperita
; IMMEDIATE SOURCE:
; CLONE: pPM17
; US-08-881-784-8

Query Match 14.0%; Score 242.2; DB 3; Length 1665;
Best Local Similarity 50.1%; Pred. No. 2,9e-55;
Matches 691; Conservative 0; Mismatches 673; Indels 15; Gaps 3;
QY 199 CTTGGCTCAAAATTTATGACCTGTTATGAGTCTTCAAATTTGCCAAAGTTTCAGCTGTTGT 258
Db 204 CGTGGCGAAGCAGTACGGCCCACTGCGCAGCTGCGAGGTGTTCTCCGTCT 263
QY 259 CATTTCCTTCAGCTGAAGCAGCCAGAGAGGTTATGAAACTCAGGCTGATGCCCTTCGCCCA 318
Db 264 GCTCTGCTCCCGAGGCCACCAAGAGGCGATGAAGCTGGTGGACCCCGCTTCGCGCGA 323
QY 319 AGCGCTTATGCTTTGGACGACAGATTGTTGTTTTTATATCGAAAGATGTTCTGTTTTCG 378
Db 324 CCGGTTTCAGAGCATCGGAGCGAAGATCATGTGTACGACACGACGACATCATCTTCAG 383
QY 379 TTCATATGAGATCATCTGAGGACAGATGAAGAAAATTTGGATCTTTGAATTTCTGAGTGC 438
Db 384 CCCCTACAGCGTGCACTGGCGCCAGATCGGAAGATCTGCGTCTCCGAGCTCTCAGCGC 443
QY 439 CAAAAAAGTTCAATCCTCCAGGTTAATCCGAGAGAGAAATGGAGGATGCCATCACATT 498
Db 444 CCGCAACGTCCGCTCTTCGGCTTCATCAGGACGAGACGAGGTGTCCTCCGCTCTCTCGGCCA 503
QY 499 CTTCCGTTTCGAAAGCC-----GGATCTCCGGTCAATATTACAAAGATCATTTATGGCAT 552
Db 504 CTTCCGCTCTCCGCGCGGGGGGAGGCCGTGACCTCACGGAGCGGATACGACGCT 563
QY 553 TATAATTTTCGATCATGATAAGAACATCCGTTGGTGAATTTGTAA---GCAAAAAGAAAGATT 609
Db 564 GACGTGCTCCATCATCTGCGAGGGCGGCTTCGGGAGCGGTGATCAGGGACACGAGAGCT 623
QY 610 GCTGAGTGTTCGGATGCAATGAGGCGACGAGTTTGGCCACCAGAGCGCTTT 669
Db 624 GGTGAGCTGTTGAAGGACGCCCTCAGCATGGCGTCCGGGTTTCGAGCTCGCCGACATGTT 683
QY 670 TCCGAGTGGAAATTTACTTTCATATATCTGAGCTGAGTCAAAACCCAGCGTTTGA 729
Db 684 CCCCTCTCCAAAGCTCCTCACTTCTGCTGGAACAAGAGCAAGCTGTGGAGGATGCG 743
QY 730 TCAGGAGATTGACGATATACCTTGAAGAGATTTTAAATGAACAAAGCCCAATAAG----- 784
Db 744 CCGCGCGCTGACGCGCATCTCGAGGCCATCTGAGGAGGACCAAGCTCAAGAAGAGCGG 803
QY 785 -CCTTTTGAAGCGGATACTTAATGATGTTCTATTGATCTTCAAAAAATGGAAACGT 843
Db 804 CGAGTTTGGCGGCGAGGACATTAATTGACGTACTCTTTAGGATGCAAGAGGATAGCCAGAT 863
QY 844 TCCAGTGCAGTGAACAAAGCAAGCATCAAAAGCATCCGTTTTCGAAAATGTTTACTGCCCG 903
Db 864 CAAAGTCCCATCACCACCAAGCCATCAAGCCCTTCATCTTCGACAGCTTCTCAGCGGG 923
QY 904 GAGCGAAACAACTTCGAAAGCTACAGATGGGTAAATGGCAGAGCTGATGAAAAATCCAAC 963
Db 924 GACCGAGACATCATCAACCAACCCCTGTGGGTGATGGCGGAGCTGATGAGGAATCCAGA 983

964	Qy	TGAACTAAGAAAGACACAAGAAGAGTTAGACAAGTAGTATTTGGTGAATGGGAAAAGTTGA	1023
984	Db	GGTGATGGCAAAACGCGAGGCGAGGTGAGAGCGCGCTGAAGGGGAAGACGGACTGGGA	1043
1024	Qy	TGAATCAAGATTTCATGATTTGAAATTCITCAAGTTAGTGGTTAAAGAAAACCTTAAGATT	1083
1044	Db	CGTGGACGACGTGCAGGAGCTTAAGTACATGAAATCGTGGTGAAGGAGACGATAGGAT	1103
1084	Qy	ACATCCTCCCGTGTCTTTGATTCGAGGGAGTGTAGAGAAAACAACGAAATTCGATGGATA	1143
1104	Db	GCACCTCCGATCCGTTGATCCCGAGATCATGCGAGAGAGAAATGCGAGGTCAACGGGTA	1163
1144	Qy	TGAAATTCATCCGAACACTCGAATTTGTTGAAATGCTTGGCGGATAGGAAGAGATCCTAA	1203
1164	Db	CACGATTTCCGAATAAGGCCAGAAATCATGATCAACGTGTGTGCCATGGTAGGAATCCTCT	1223
1204	Qy	TACTTGGTCGGAACCTCGGAAAGTTTAACCCAGAAAGGTTTTAAAGATTGTGCAATTGATTA	1263
1224	Db	CTACTGGGAAAAACCCGAGACCTTTTGGCCGAAAGGTTTTGACCAAGCTCTCGAGGGATTT	1283
1264	Qy	TAAAGGGACGACATTTGCACTGGTGACATTTGGTGCAAGGAAAAGAATATGTCTCTGGCAT	1323
1284	Db	CATGGGAAAACGATTTCGAGTTTCATCCATTTGGAGCTTGGAAAGAAGAATCTGCCCGGTTT	1343
1324	Qy	TACTTCAGCTATTAACCAATTTGGAGTATGTCAATTATAAATCTATTATATCATTTTAATG	1383
1344	Db	GAAATTCGGGTTGGCAAAATGTTGAGGTCCCAATTGGCACAGCTTCTTTACCACTTCGACTG	1403
1384	Qy	GGAACTGGCCGATGGAAATTAACCTCAAAACATTTGATATGACTGAACTATTGGCGGTGC	1443
1404	Db	GAACTTGGCGGAAGGAATGAACCCCTCCGATATGGACATGTCTGAGGCAGAAGGCCCTTAC	1463
1444	Qy	TCTCAGGAAAAAATAAGATCTTAAGTTGATTCCTATTTCCTATATCAAGTTTAGCTTAGGCTC	1503
1464	Db	CGGAATTAAGAAGAACAATCTTCTACTGTTCCACACCCCTACGATCCTTCTCTCATGATC	1523
1504	Qy	AAATATTTCTTGATTACATAGGAGGGTTGAAATATATATAATAAACTTTTAATTAACGAT	1562
1524	Db	AATTAATACCTCTTTAAATTTGCTCCCTTTGAAATAAGAGTGCATATACATATATGATATAT	1582

RESULT 14

```

US-09-292-768-3
; Sequence 3, Application US/09292768
; Patent No. 6194185
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Lupien, Shari L
; APPLICANT: Karp, Frank
; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR THE PRODUCTION OF
; TITLE OF INVENTION: LIMONENE HYDROXYLASES
; FILE REFERENCE: wbur13463
; CURRENT APPLICATION NUMBER: US/09/292,768
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 08/881,784
; EARLIER FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: Mentha piperita
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19)..(1518)
US-09-292-768-3

```

Query Match 14.0%; Score 242.2; DB 3; Length 1665;
Best Local Similarity 50.1%; Pred. No. 2.9e-55;
Matches 691; Conservative 0; Mismatches 673; Indels 15; Gaps 3;
Ov 199 CTTGGCTCAAAATTTATGACACTGTTATGATGCTTCAAATTTGGCCAAGTTTCAGCTGTGT 258

Db 1284 CATGGGAACGATTCGAGTTCATCCATTTGGAGCTGGAAGAAGAAATCTGCCCGGTTT 1343
QY 1324 TACTTCAGCTATTACCAATTTGGAGTATGTCATTATAAATCTATTATATCATTTTAATG 1383
Db 1344 GAATTTCCGGTTGGCAATGTTGAGTCCATTTGGCACAGCTTCTTTACCACTTCAGCTG 1403
QY 1384 GGAACCTGGCGATGGAATTAACCTCAACCTTGAATGATGATGATGATGATGATGATGATG 1443
Db 1404 GAAGTTGGCGGAAGGAATGAACCTTCCGATATGGACATGTCGTGAGCAGAAAGCCCTTAC 1463
QY 1444 TCTCAGGAAGAAATAGATCTTAAGTGTGATTCCTATTCCATATCAAGTTAGCTTGGCTC 1503
Db 1464 CGGAATAAGAAAGAAATCTTCTACTCGTTCCACACCTACGATCCTTCTCTCATGATC 1523
QY 1504 AAATATTCTTGTATACATAGGAGGTGGAATATATATATAATAAATCTTTAAATTAACGAT 1562
Db 1524 AATTAATCTCTTAATTTGCTCTTGAATAAAGAGTGCATATATATATATATATATATAT 1582

RESULT 15
US-09-292-768-69
; Sequence 69, Application US/09292768
; Patent No. 6194185
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Lupien, Shari L
; APPLICANT: Karp, Frank
; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR THE PRODUCTION OF
; TITLE OF INVENTION: LIMONENE HYDROXYLASES
; FILE REFERENCE: wsur13463
; CURRENT APPLICATION NUMBER: US/09/292,768
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 08/881,784
; EARLIER FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 69
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: computer-generated nucleic acid sequence encoding
; OTHER INFORMATION: limonene-3-hydroxylase
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(1665)
; OTHER INFORMATION: computer-generated nucleic acid sequence encoding
; OTHER INFORMATION: limonene-3-hydroxylase protein variant
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19)..(1518)
US-09-292-768-69

Query Match 14.0%; Score 242.2; DB 3; Length 1665;
Best Local Similarity 50.1%; Pred. No. 2,9e-55;
Matches 691; Conservative 0; Mismatches 673; Indels 15; Gaps 3;

QY 199 CTTGGCTCAATTTATGACCTGTTATGATCTTCAATTTGGCCAAAGTTTCAGCTGTTGT 258
Db 204 CGTGGCAAGCAGTACGGCCCAAGTGGCGCACGTCGACGCTCGGCGAGGTCTTCCGTCGT 263
QY 259 CATTTCTTCAGCTGAAGCAGCCAAAGAGGTATGAAAATCTCAGGCTGATGCCCTTCGCCCA 318
Db 264 GCTCTCGTCCCGAGGCCACGAAGAGCGCATGAGCTGGGACCCGGCCCTGCGCGGA 323
QY 319 AGCCCTCATGCTTTGGACGACAGATTGTTGTTTATATATCGAAAGAGTGTCTTGTTCG 378
Db 324 CCGGTTCCGAGACATCGGACCAAGATCATGTGTTGACGACACGACGACATCATCTTCAG 383
QY 379 TTCATATGAGATCACTGAGGAGCAGATGAAGAAAATTTGGATACTTTGAATTTCTGAGTGC 438
Db 384 CCCCTACAGCGTGCATCTGGCGCCAGATGCGGAAGATCTGCGTCTCCGAGCTCTCAGCGC 443

QY 439 CAABAAGTTCAATCCTCCAGCTTAATCCGAGAGGAGAAATGGAGGATGCCATCACATT 498
Db 444 CGCAACGTCGCTCTTTTCGGCTTTCATCAGGAGGAGGAGGTGTCCTCGGCTCTCTCGGCCA 503
QY 499 CTTCCGTTTCAAAAGCC-----GGATCTCCGGTCAATATTACAAAGATCATTTTATGGCAT 552
Db 504 CTTCCGCTCTCTCGGCGCGCGGGGAGGCGCTGGACCTCACGAGCGGATACGAGCGCT 563
QY 553 TATATATTTTCGATCATGATAAGAACATCCGTTGGTATTTGTAA---GCAAAAAAGAAAGATT 609
Db 564 GAGTGTCTCATCATCTGAGGCGCGCTTCGAGGAGCTGATCAGGAGACCAACGAGAGCT 623
QY 610 GCTGAGTGTTCGCGATGACAGTCAATGAGGACGAGCTGTTTGGCACCGCACGCTTT 669
Db 624 GGTGAGCTGTTGAAGGACGCTCCTGAGTGGCGTTCGGGTTTCGAGCTCGCCACATGTT 683
QY 670 TCCGAGCTGGAATTAATTTCACTATATATCATTTGAGCTGAGTCAAAACCCAGGCGTTTGCA 729
Db 684 CCCCTCTCCAAGCTCTCAACTTGTCTGCTGGAACAGAGCAAGCTGTGGAGGATGCG 743
QY 730 TCAGGAGATTGACGATATCTTTGAAGAGATTCTTAATGAACACAAAGCCAAATAG----- 784
Db 744 CCGCGCGCTCGACGCGCATCTCGAGGCCATCTCGAGGCCATCTGAGGAGCAAGCTCAAGAAGAGCGG 803
QY 785 -CCTTTTGAAGCGGATACTTAATGAGTGTCTTATTTGATCTTCAAAAAATCGAAACCT 843
Db 804 CGAGTTTGGCGGAGGACATTAATGACGTACTCTTTAGGATGCAAGAGGATAGCCAGAT 863
QY 844 TCCAGTGCAGTGACAAACGAAAGCATCAAGCATCCGTTTTCGAAATGTTTACTGCCCG 903
Db 864 CAAAGTCCCATCACCAACGCGCATCAAGCCTTCATCTTCGACACGCTCTCAGCGGG 923
QY 904 GAGCGAAACAACTTCGAAAGCTACAGATGGGTATGCGAGAGCTGATGAAAAATCAAC 963
Db 924 GACCGAGACATCATCAACCCACCTGTGGGTGATGGCGGAGCTGATGAGGAATCCAGA 983
QY 964 TGAACTAAGAAAACACAGAAAGATTAGACAAGTATTGTTGTAATCGGAAAAGTTGA 1023
Db 984 GGTGATGGGAAAGCGCAGCGAGGTGAGACGCGGCTGAAGGGGAAACGAGCTGGGA 1043
QY 1024 TGAATCAAGATTTTCATGATTTGAAATCTTCAAGTTAGTGGTTAAAGAAACTCTAAGATT 1083
Db 1044 CGTGACGACGTCGAGGAGCTTAAGTACATGAAATCGGTGTTGAGGAGACGATGAGGAT 1103
QY 1084 ACATCTCTCGGTTGCTTTGATTCGAGGAGGTGTAGAGAAACAAACAGAAATGATGGATA 1143
Db 1104 GCACCTCCGATCCCGTTGATCCCGAGATCATGCGAGAGAAATGCGAGGTCAACGGGTA 1163
QY 1144 TGAAATTCATCGGAACACTCGAATTTGTTGTAATCGTTGGCGATAGGAAGAGATCCTAA 1203
Db 1164 CACGATTCGAAATAAGCCAGAAATCATGATCAACGTGTGTCATGCGGTAGGAATCCTCT 1223
QY 1204 TACTTGGTCGGAACCTGGAAAAGTTTAAACCCAGAAAAGTTTAAAGATTTGTGCAATTTGATTA 1263
Db 1224 CTACTGGGAAAACCCCGAGACCTTTTGGCCGAAAGTTTGACCAAGTCTCGAGGATTT 1283
QY 1264 TAAAGGAGCAGATTTTGAATCTGATACCATTTGGTCCAGAAAAGAAATATGTCCTGGCAT 1323
Db 1284 CATGGAAAACGATTTTCGAGTTTCATCCCATTTGGAGCTCGAAGAAAGAAATCGCCCGGTTT 1343
QY 1324 TACTTCAGCTATTACCAATTTGGAGTATGTCATTATAATCTATTATATCATTTTAAATG 1383
Db 1344 GAATTTCCGGTTGGCAATTTGTTAGGTCCTTGGCACAGCTCTTTTACCATCTTGACTG 1403
QY 1384 GGAATGCGCGATGGAATTTACACCTCAAAACACTTGTATATGACTGAAGCTATTGGCGGTGC 1443
Db 1404 GAAAGTTGCGGAAGAAATGAACCTTCCGATATGACATGTCTGAGGAGAGAGGCTTAC 1463
QY 1444 TCTCAGGAAAATAATAGATCTTAAGTGTATCTTATTCATATCAAGTTAGTTAGTCTC 1503
Db 1464 CGGAATAAGAAAGAAACAAATCTTCTACTCGTTCCCAACCCCTACGATCCTCTCTATGATC 1523

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 31, 2006, 06:03:08 ; Search time 8489 Seconds
(without alignments)
5649.731 Million cell updates/sec

Title: US-10-759-813-2

Perfect score: 2574

Sequence: 1 MEQKLSFPISILISFLVLI.....KIDKLIPYQVSLGNSIS 500

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-Q=/abs/ABSSWEB_spool/US10759813/runat_31052006_060304_2592/app_query.fasta_1
-DB=GenEmbl -QFMT=fastap -SUFFIX=rg -MINMATCH=0.1 -LOOPECL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs05h
-USER=US10759813 @CN_1_1_5548@runat_31052006_060304_2592 -NCPU=6 -ICPU=3
-NO_MMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*

1: gb_env.*
2: gb_pat.*
3: gb_ph.*
4: gb_pl.*
5: gb_pr.*
6: gb_ro.*
7: gb_sta.*
8: gb_sy.*
9: gb_un.*
10: gb_vl.*
11: gb_ov.*
12: gb_hcg.*
13: gb_in.*
14: gb_om.*
15: gb_ba.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Match Length	Description
1	2574	100.0	AX366075 Sequence
2	2574	100.0	AF406732 Euphorbia
3	1164	45.2	CS223906 Sequence

4	1149	44.6	1673	2	CS223964	Sequence
5	1149	44.6	1673	2	CS223984	Sequence
6	1148	44.6	1673	2	CS223962	Sequence
7	1144	44.4	1657	2	CS223948	Sequence
8	1142	44.4	1673	2	CS223966	Sequence
9	1142	44.4	1701	4	STU296346	Solanum t
10	1131	43.9	1610	2	CS223950	Sequence
11	1131	43.9	1610	2	CS224014	Sequence
12	1129.5	43.9	1576	2	CS223902	Sequence
13	1128.5	43.8	1566	2	CS223944	Sequence
14	1128.5	43.8	1581	2	CS223904	Sequence
15	1128.5	43.8	1664	2	CS223952	Sequence
16	1127.5	43.8	1781	2	E63735	Cytochrome
17	1127.5	43.8	1781	2	AR205821	Sequence
18	1127.5	43.8	1781	2	AF122821	Capsicum
19	1122.5	43.6	1784	2	CS249182	Sequence
20	1119	43.5	1665	2	A29004	Sequence
21	1119	43.5	1665	2	I47318	Sequence 29
22	1106.5	43.0	1509	4	AB028462	Petunia x
23	1102	42.8	99089	4	AP007265	Lotus cor
24	1100	42.7	1515	4	AF368376	Nicotiana
25	1097.5	42.6	1572	2	CS223930	Sequence
26	1088	42.3	1666	2	I28313	Sequence 24
27	1088	42.3	1666	2	CS224040	Sequence
28	1082.5	42.1	1559	2	GMC450CP7	mRNA
29	1071.5	41.6	1800	4	CR954196	Medicago
30	1068	41.5	132419	4	GMC450CP3	
31	1059.5	41.2	1754	4	DQ335785	Medicago
32	1054.5	41.0	1789	4	AC140774	Medicago
33	1054.5	41.0	116998	4	AF000403	Solanum cha
34	1052.5	40.9	1641	4	SCU48434	
35	1045.5	40.6	1641	4	AF000403	Lotus jap
36	1042.5	40.5	1682	4	CR0238612	Catharan
37	1039.5	40.4	1506	4	AY596975	Centauri
38	1036	40.2	3107	4	SCU48435	
39	1026.5	39.9	122392	4	AC137079	Solanum cha
40	1026	39.9	110957	12	CT009652	Medicago
41	1026	39.9	122965	4	CR932040	Medicago
42	1023	39.7	139522	4	AC145061	Medicago
43	1023	39.7	255855	12	AC145024	Medicago
44	1022.5	39.7	1564	2	CS223914	Sequence
45	1019	39.6	122965	4	CR932040	Medicago

ALIGNMENTS

RESULT 1	AX366075	Sequence 1	1733 bp	DNA	linear	PAT 15-FEB-2002
LOCUS	AX366075	Sequence 1	from Patent WO208269.			
DEFINITION	AX366075					
ACCESSION	AX366075					
VERSION	AX366075.1	GI:18697502				
KEYWORDS						
SOURCE	Euphorbia lagascae					
ORGANISM	Euphorbia lagascae					
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;					
	rosids; eurosids I; Malpighiales; Euphorbiales; Euphorbiaceae; Euphorbioideae;					
	Euphorbiae; Euphorbia.					
REFERENCE	1					
AUTHORS	Cahoon, E.B.					
TITLE	A cytochrome p450 enzyme associated with the synthesis of _g(d)?12					
	-epoxy groups in fatty acids of plants					
JOURNAL	Patent: WO 0208269-A 1 31-JAN-2002;					
	E. I. du Pont de Nemours and Company (US)					
FEATURES	Location/Qualifiers					
	1..1733					
	/organism="Euphorbia lagascae"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:54672"					

ORIGIN

Alignment Scores:

```
Pred. No.: 1,11e-286 Length: 1733
Score: 2574.00 Matches: 500
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-759-813-2 (1-500) x AX366075 (1-1733)

QY 1 MetGluGlnLysAsnLeuSerPheProSerIleLeuIleSerPheLeuLeuValLeuIle 20
DB 14 ATGGAGCAGAAAAATCTCTCTTTCCGAGCAATTTTAATAAGTTTCTGCTGTTTAAATC 73
QY 21 LeuValValMetArgLeuTrpLysLysGlnAsnProProProGlyProTrpLysPhe 40
DB 74 TTAGTAGTAGTCATGAGGTGTGGAGAAACAGAAATCCACCTCCAGGGCCATGGAAGTTT 133
QY 41 ProIleIleGlyAsnLeuProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPhe 60
DB 134 CCTATCATAGGTAAATCTCTCATTTATTTACTACCTTCGTGATCTAGGCCCATGAACTTTT 193
QY 61 ArgAlaLeuAlaGlnIleTrpGlyProValMetSerLeuGlnIleGlyGlnValSerAla 80
DB 194 AGAGCCTTGGCTCAAAATTTATGGACCTGTTATGAGTCTTCAAAATTTGGCCAGTTTCAGCT 253
QY 81 ValValIleSerSerAlaGluAlaLysGluValMetLysThrGlnAlaAspAlaPhe 100
DB 254 GTTGTCAATTTCTTCACTGAAGCAGCCAAAGAGGTTATGAAACTCAGGCTGATGCCCTTC 313
QY 101 AlaGlnArgProIleValLeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeu 120
DB 314 GCCCAACGCCCTATGCTCTTGACGACACAGATTGTTTATATCGGAAAGATGCTCTG 373
QY 121 PheAlaSerTyrGlyAspHisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeu 140
DB 374 TTTTGTTCATATGGAGATCACTGGAGGCAGATGAGAAATTTGGATACCTTCAATTTCTG 433
QY 141 SerAlaLysLysValGlnSerSerArgLeuIleArgGluGluMetGluAspAlaIle 160
DB 434 AGTGCCAAAAAAGTTTCAATCTCTCAGGTTAATCCGAGAGAGAGAAATGGAGGATGCCATC 493
QY 161 ThrPheLeuArgSerLysAlaGlySerProValAsnIleThrLysIleIleTrpGlyIle 180
DB 494 ACATTCCTCCGTTGAAACCCGATCTCCGGTCAATATTACAAAGATCATTTATGGCATT 553
QY 181 IleIleSerIleMetIleArgThrSerValGlyAsnCysLysGlnLysGluArgLeuLeu 200
DB 554 ATAATTTTCGATCATGATAGAACATCCGTTGGTTGTTAATGTAACAAAAAGAAAGATTGCTG 613
QY 201 SerValAlaAspAlaValAsnGluAlaIleThrSerPheGlyThrAlaAspAlaPhePro 220
DB 614 AGTGTTCGCCGATGCAGTCAATGAGCGAGCGAGGTTTTTGGCACCAGCAGACCTTTTCGG 673
QY 221 ThrTrpLysLeuLeuHisTyrIleIleGlyAlaGluSerLysProArgArgLeuHisGln 240
DB 674 AGTGGAATATTCTTCACTATATCATTTGGAGCTGAGTCAAAACCCAGGGGTTTGATCAG 733
QY 241 GluIleAspAspIleLeuGluIleLeuAsnGluHisLysAlaAsnLysProPheGlu 260
DB 734 GAGATTGACGATATATCTTGAAGAGATTCTTAATGAACACAAAGCCCAATAAGCCTTTTGA 793
QY 261 AlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValProValPro 280
DB 794 GCGGATAACTTAATGATGTTCTATTGAAATCTTCAAAAAAATGGAAACGTTCCAGTGCCA 853
QY 281 ValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySerGluThr 300
DB 854 GTGACAAAAGAAAGCATCAAGCATCCGTTTTTGCAATGTTTACTGCGGGAGGAGAAACA 913
QY 301 ThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGluLeuArg 320
DB 914 ACTTCGAAAGCTACAGATGGGTAAATGGCAGAGCTGATGAAAAATCCAATCACTGAACTAAGA 973
```

```
QY 321 LysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGluSerArg 340
DB 974 AAAGCACAGAAGAAGATTAGACAAGATATTTGGTGAATGGGAAAGTTGTGTAATCAAGA 1033
QY 341 PheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHisProPro 360
DB 1034 TTTTCATGATTTGAAATCTTCAAGTTAGTGTGTTAAAGAAACTCTAAGATTACATCCTCG 1093
QY 361 ValValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTrpGluIleHis 380
DB 1094 GTTGTCTTGTATCCGAGGAGTGTAGAGAAACAAACACGAATTTGATGGATATGAAATTCAT 1153
QY 381 ProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsnThrTrpSer 400
DB 1154 CCGAACACTCGAATTTGTTGAATGCTTGGGGATAGGAAGATCCTAATACTTGGTTCG 1213
QY 401 GluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyrLysGlyThr 420
DB 1214 GAACCTGGAAGTTTAAACCCAGAAAGTTTAAAGATTGTGCAATTTGATTATAAGGACG 1273
QY 421 ThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIleThrSerAla 440
DB 1274 ACATTTGAACCTGGTACCATTTGGTGCAGGAAAAAGAAATATGTCTGGCATTTACTTCAGT 1333
QY 441 IleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrpGluLeuAla 460
DB 1334 ATTACCAATTTGGAGTATGTCATTATAAATCTATTATATCATTTTAAATTTGGAACTGGCC 1393
QY 461 AspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAlaLeuArgLys 480
DB 1394 GATGGAATTCACCTCAAAACACTTGATGATGACTGAAGCTATTGGCGGTGCTCTCAGGAAA 1453
QY 481 LysIleAspLeuLysLeuIleProIleProTyrGlnValSerLeuGlySerAsnIleSer 500
DB 1454 AAAATAGATCTTAAAGTTGATTTCTTATCCATATCAAGTTAGCTTAGGCTCAAATATTTCT 1513

RESULT 2
AF406732 1758 bp mRNA linear PLN 13-FEB-2002
LOCUS Euphorbia lagascae cytochrome P450 (CYP726A1) mRNA, complete cds.
DEFINITION AF406732
ACCESSION AF406732
VERSION AF406732.1 GI:18157658
KEYWORDS
SOURCE Euphorbia lagascae
ORGANISM Euphorbia lagascae
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Euphorbiaceae; Euphorbioideae;
Euphorbieae; Euphorbia.
1 (bases 1 to 1758)
Cahoon,E.B., Ripp,K.G., Hall,S.E. and McGonigle,B.
Transgenic production of epoxy fatty acids by expression of a
cytochrome P450 enzyme from Euphorbia lagascae seed
Plant Physiol. 128 (2), 615-624 (2002)
11842164
2 (bases 1 to 1758)
Cahoon,E.B.
Direct Submission
Submitted (06-AUG-2001) DuPont Crop Genetics, E.I. DuPont Company,
E402/4212, Wilmington, DE 19880-0402, USA
FEATURES
source
1..1758
/organism="Euphorbia lagascae"
/mol_type="mRNA"
/db_xref="taxon:54672"
1..1758
/gene="CYP726A1"
21..1523
/gene="CYP726A1"
/feature="delta2-fatty acid epoxigenase"
/codon_start=1
/product="cytochrome P450"
/protein_id="NAL62063.1"
```

/db_xref="GI:18157659"
/translation="MEQKNLSPPSILISFLVLLVVMRLWKKQNPFPKPKFIIG
NLPHLTSLDLGHERFRAAOIYGPVMSIQIGOVSAVWISSAEAAKEVNMKTOADAPAO
RPIVLDAQOIVFYNRKDVLFSYGDHWRQMKIWIIEFLSAKKVQSSRLUREMEDAI
TFRLSAGSVNITKIIYGIISIMIRTSVGNCKQKERLGSVADAVNEAATSGTADA
FPTWKLHYSIIAGESPRLHQEIDIDILEILNEHKANKPFADNLMDVLLNLQKNG
VFPVTNESIKASVLQMFAGSETTSKATEWMAELMKNPTELKRAQEVRQVFGEMG
KYDESFDLKFELVVKETLRLHPVLIIPRECRTTIDGVEIHPNTRIIVNNAWI
GRDNTWSPPGFNPRFKCAIDYKGTTFELVPPFGAGKRIQPGITSALTINLEYVIN
LLYHFNWELADGITPQTLDMTEAIGGALKKIDLKLIPITYQVSLGNSIS"
ORIGIN

Alignment Scores:
Pred. No.: 1,13e-286 Length: 1758
Score: 2574.00 Matches: 500
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 4 Gaps: 0
US-10-759-813-2 (1-500) x AF406732 (1-1758)

Qy	1	MetGluGlnLysAsnLeuSerPheProSerIleLeuIleSerPheLeuLeuValLeuIle	20
Db	21	ATGGAGCAGAAAATCTCTTTCCGAGCATTTTAATAAGTTTTCGCTGTGTTTAAATC	80
Qy	21	LeuValValValMetArgLeuTrpLysLysGlnAsnProProGlyProTrpLysPhe	40
Db	81	TTAGTAGTAGTCATGAGGTGTGGAAGAAACAGATCCACCTCCAGGGCCATGGAAGTTT	140
Qy	41	ProIleIleGlyAsnLeuProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPhe	60
Db	141	CCTATCATAGGTAATCTTCTCATTTATTACTCACCTTCATGATAGGCCATGAACGTTT	200
Qy	61	ArgAlaLeuAlaGlnIleTyrGlyProValMetSerLeuGlnIleGlyGlnValSerAla	80
Db	201	AGAGCCTTCGGCTCAAAATTTATGGACCTGTTATGAGTCTTCAAAATGGCCAAAGTTT	260
Qy	81	ValValIleSerSerAlaGluAlaLysGluValMetLysThrGlnAlaAspAlaPhe	100
Db	261	GTGTGTCATTTCTTACGTGAAGCAGCCAAAGAGGTTATGAAAACCTCAGGCTGATGCC	320
Qy	101	AlaGlnArgProIleValLeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeu	120
Db	321	GCCCAAGCCCTATCGCTTTCGAGCGCACAGATTTGCTTTTATAATCGGAAGATGCTTG	380
Qy	121	PheAlaSerTyrGlyAspHisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeu	140
Db	381	TTTGCTTCATATGGAGATCACTGGAGGCAGATGAAAGAAAATTTGGATCTTGAATTTCTG	440
Qy	141	SerAlaLysLysValGlnSerArgLeuIleArgGluGluMetGluAspAlaIle	160
Db	441	AGTGCCAAAAGATTCATCTCCAGGTTAATCCGAGAGGAAGAAATGGAGAGTCCATC	500
Qy	161	ThrPheLeuArgSerLysAlaGlySerProValAsnIleThrLysIleIleTyrGlyIle	180
Db	501	ACATTCTCCGTTTGGAAACCGGATCTCCGGTCAATATTACAAAGATCATTTATGCAATT	560
Qy	181	IleIleSerIleMetIleArgThrSerValGlyAsnCysLysGlnLysGluArgLeuLeu	200
Db	561	ATPAATTCGATCATGATAGAACATCCGTTGGTAATTGTAAGCAAAAGAAAGATTGCTG	620
Qy	201	SerValAlaAspAlaValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPhePro	220
Db	621	AGTGTGCCGATCGAGTCAATGAGGCGCGACGAGTTTGGCACCCGACACGCTTTTCGG	680
Qy	221	ThrTrpLysLeuLeuHisTyrIleIleGlyAlaGluSerLysPheProArgArgLeuHisGln	240
Db	681	ACGTGGAATTAATCTTACCTATATCATTTGAGCTGAGTCAAAACCCAGCGGCTTTCATCAG	740
Qy	241	GluIleAspAspIleLeuGluGluLeuAsnGluHisLysAlaAsnLysProPheGlu	260
Db	741	GAGATTGACGATATATCTTGAAGAGATTCTTAATGAACACAAAGCCCAATAAGCCTTTTGA	800

Qy	261	AlaAspAsnLeuMetAspValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu	280
Db	801	GCGGATAACTTAATGATGTTCTTATTGAATCTTCAAAAATAATGGAACGTTCCAGTGCCA	860
Qy	281	ValThrAsnGluSerIleLysAlaSerValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu	300
Db	861	GTGACAAACGAAAGCATCAAGCATCCGTTTGGCAATGTTTACTCTCCGGAGCGGAAACA	920
Qy	301	ThrSerLysAlaThrGluTrpValMetAlaGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu	320
Db	921	ACTTCGAAAGCTACAGAATGGGTAAATGGCAGAGCTGATGAAAAATCCAACGTGAACTAAGA	980
Qy	321	LysAlaGlnGluGluValArgGlnValPheGlyGluMetGlyLysValAspGluSerArg	340
Db	981	AAAGCACAGAGAGAGTAGACAGATTTTGGTGAATAATGGGAAAAAGTTGATGATCAAGA	1040
Qy	341	PheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHisProPro	360
Db	1041	TTTTCATGATTGAAATTTCTCAAGTTAGTGGTTAAAGAAACTCTAAGATTTACATCTCCG	1100
Qy	361	ValValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyrGluIleHis	380
Db	1101	GTGTCTTGTGATTCGAGGGAGTGTAGAGAAACAACAAGATTCATGATGATGAATTCAT	1160
Qy	381	ProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsnThrTrpSer	400
Db	1161	CCGACACCTCGAATTTGTTGTAATGCTTGGGCGATAGGAAGAGATCTTAATCTTGGTGC	1220
Qy	401	GluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyrLysGlyThr	420
Db	1221	GAACCTGGAAAGTTTAAACCCAGAAAGTTTAAAGATTTGCAATTCATTAATAAGGACG	1280
Qy	421	ThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIleThrSerAla	440
Db	1281	ACATTGAACCTGATCCATTTGCTGCAGGAAAAAGAAATATGCTTGGCCTTACTTCAGCT	1340
Qy	441	IleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrpGluLeuAla	460
Db	1341	ATTACCAATTTGGAGTAGTGTCAATTAATCATTTATATCATTTTAATTTGGGAAC TGCC	1400
Qy	461	AspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAlaLeuArgLys	480
Db	1401	GATGGAATTTACACCTCAACACTTGTATGACTGAAGCTATTGGCGGTGCTCTCAGGAAA	1460
Qy	481	LysIleAspLeuLysLeuIleProIleProTyrGlnValSerLeuGlySerAsnIleSer	500
Db	1461	AAAAATAGATCTTAAGTTGATTCCTATTCCATATCAAGTTAGCTTAGGCTCAAAATATTC	1520
RESULT 3			
LOCUS	CS223906	1538 bp	DNA linear PAT 15-DEC-2005
DEFINITION	Sequence 162 from Patent WO2005111217.		
ACCESSION	CS223906		
VERSION	CS223906.1		
KEYWORDS	GI:83685434		
SOURCE	Nicotiana tabacum (common tobacco)		
ORGANISM	Nicotiana tabacum		
REFERENCE	1		
AUTHORS	Xu D.		
TITLE	Nicotiana nucleic acid molecules and uses thereof		
JOURNAL	Patent: WO 2005111217-A 162 24-NOV-2005;		
FEATURES	U.S. Smokeless Tobacco Company (US)		
	Location/Qualifiers		
	1..1538		
source	/organism="Nicotiana tabacum"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:4097"		
ORIGIN			

Db 54 TTGGTTTCCATTTCCTATTCTATCTTCCTCTCTTTTGTAAAGATATGGAAGAACTCC 113
QY 32 AnPro-----ProProGlyProTrpLysPheProIleIleGlyAsnLeu 46
Db 114 NATAGCAAGCAAAAGTTGCCACGAGTCCATGGAACTACCAATACTAGGAAGTAG 173
QY 47 ProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle 66
Db 174 CTTTCATATGGTT-----GGTGGACTACCACACCATGTCTCTTAGAGATTTAGCCAAAAA 227
QY 67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSerAla 86
Db 228 TATGGACCACTTATGACCTTCAATAGGTGAAGTTTCTGCGGTTCGTCTCTCTCT 287
QY 87 GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
Db 288 GATACGGCAAAAGAGATTAATAAACTCATGACATCGCTTTTCGGTCTAGGCTACGCTT 347
QY 107 LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
Db 348 TTGGCCCCGAGATTGTCTGTACAAATAGTCTGTATAGCTTTTGGCCCTATGGCGAC 407
QY 127 HisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln 146
Db 408 TATTGGAGACAAATGCGTAAATATATGTCTTGGAAAGTGTCTGAGTCCCAAGAATGTTCG 467
QY 147 SerSerArgLeuLeuArgGluGluGluMetGluAspAlaIleThrPheLeuArgSerLys 166
Db 468 ACATTAGTCTTATTAGGCGGAATGAAGTTCTTGGTCTCATTAATTTATCGGCTCATCT 527
QY 167 AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleIleSerIleMetIle 186
Db 528 TCTGGTGAACCTATTAAATGTACGGAAGGATCTTTTGTTCACAAGCTCCATGACATGT 587
QY 187 ArgThrSerValGlyAsnCys---LysGlnLysGluArgLeuLeuSerValAlaAspAla 205
Db 588 AGATACGCGTTTGGGCAAGTGTTCAAAGAGACAGCAAAATTTATCAACTAATTAAGAA 647
QY 206 ValAsnGluAlaIleThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu 225
Db 648 GTGATACTTTAGCAGGAGGTTTGATGTGGCTGCATATTCCTTCATCTGAAGTTCTT 707
QY 226 HisTyrIleIleGlyAlaGlySerLysProArgArgLeuHisGlnGluIleAspIle 245
Db 708 CATGTGCTCAGTGGGAATGAAGGTGAAGATTATGAATGCACACCAATAGTAGTCCATT 767
QY 246 LeuGluGluIleLeuAsnGluHis-----LysAlaAsnLysPro 258
Db 768 GTTGAGAATGTCAATCAATGAGCACAAAGAAAAATCTTGCATTTGGGAAAACTAATGGAGCG 827
QY 259 PheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValPro 278
Db 828 TTAGGAGGTGAAGATTAAATGATGTCTTCTTAAGACTTATGAATGATGGAGCCCTTCAA 887
QY 279 ValProValThrAsnGlnSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer 298
Db 888 TTTCTCTATCACCACGACCAACATCAAGGCCATAATTTTTCACATGTTTCTGCTCGCGGACA 947
QY 299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
Db 948 GAGACTTTCATCTCAACCAATTTGTGGGCTATGGTGAAGAAATGGTGAAGAAATCCAGCCGTA 1007
QY 319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
Db 1008 TTCGGGAAGCTCAGACAGAGTAGAGAGGACATTTAGAGGAAAAAGAACTTTTCGATGAA 1067
QY 339 SerArgPheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHis 358
Db 1068 AATGATGTGAGGAGCTAACTACCTAAAGTTAGTAAATAAAGAACTCTAGAGTTTCAT 1127
QY 359 ProProVal---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyr 377
Db 1128 CCACCGGTTCCACTTTTGTCTCCCAAGAGAATGTAGGGAAGAGACAAATATAAACCGCTAC 1187

QY 378 GluIleHisProAsnThrArgIleValValAsnAlaIleTrpAlaIleGlyArgAspProAsn 397
Db 1188 ACTATTCTCTGTAAAGACCAAGTCATGGTTAATGTTTGGCTTTGGGAAGAGATCAAAA 1247
QY 398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr 417
Db 1248 TATTGGAATGACCGAAGAACTTTATGCCAGAGAGATTTCAGCAGTGTCTTAAGGATTTT 1307
QY 418 LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIle 437
Db 1308 GTTGGTAATAATTTTGAATATATCTTCATTTGGTGGCGAAGGAGGATTTGTCTCGGAT 1367
QY 438 ThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrp 457
Db 1368 TCGTTTGGCTTAGCTAATGCTTATTTCGCATTTGGCTCAATCTATATCATCTCGATTGG 1427
QY 458 GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAla 477
Db 1428 AAATCCCTGCTGGAATCGAACCAAGGACATTGGACTTCAGTTCGTTGGTGAATACT 1487
QY 478 LeuArgLysLysIleAspLeuLysLeuIleProIleProTyrGln 492
Db 1488 GCCGCTAGAAAAAGTGACCTTTACTTGGTTGGGACTCCTTATCA 1532
RESULT 5
LOCUS CS223984 1673 bp DNA linear PAT 15-DEC-2005
DEFINITION Sequence 240 from Patent WO2005111217.
ACCESSION CS223984
VERSION CS223984.1 GI:83685871
KEYWORDS . Nicotiana tabacum (common tobacco)
SOURCE Nicotiana tabacum
ORGANISM Nicotiana tabacum
REFERENCE 1 Xu,D.
AUTHORS Nicotiana nucleic acid molecules and uses thereof
TITLE Patent: WO 200511217-A 240 24-NOV-2005;
JOURNAL U.S. Smokeless Tobacco Company (US)
FEATURES
source
1. 1673
/organism="Nicotiana tabacum"
/mol_type="unassigned DNA"
/db_xref="taxon:4097"
ORIGIN
Alignment Scores:
Pred. No.: 5,7e-122 Length: 1673
Score: 1149.00 Matches: 222
Percent Similarity: 63.2% Conservative: 91
Best Local Similarity: 44.8% Mismatches: 166
Query Match: 44.6% Indels: 16
DB: 2 Gaps: 5
US-10-759-813-2 (1-500) x CS223984 (1-1673)
QY 12 LeuIleSerPheLeuLeuValLeuIleLeuValValMetArgLeuTrpLysLysGln 31
Db 54 TTGTTTTCATTTCTTCTATCTTTCTCTTTTGTAAAGATATGGAAGAACTCC 113
QY 32 AnPro-----ProProGlyProTrpLysPheProIleIleGlyAsnLeu 46
Db 114 NATAGCAAGCAAAAGTTGCCACGAGTCCATGGAACTACCAATACTAGGAAGTAG 173
QY 47 ProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle 66
Db 174 CTTTCATATGGTT-----GGTGGACTACCACACCATGTCTCTTAGAGATTTAGCCAAAAA 227
QY 67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSerAla 86
Db 228 TATGGACCACTTATGACCTTCAATAGGTGAAGTTTCTGCGGTTCGTCTCTCTCT 287

Db	228	TATGGACCACTTATGCGACCTTCAATATAGGTGAAGTCTTCTGCGGTTGGTACTTCTCTCT	287
Qy	87	GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal	106
Db	288	GATACGGCAAAAGAAGTATTAAAACTCATGACATCGCTTTGGCTTAGCCCTAGCCCT	347
Qy	107	LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp	126
Db	348	TTGGCCCGGAGATTGTCTGTACAATAGGTCTGATCTAGCCTTTTGCCCTATGGCGAC	407
Qy	127	HisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln	146
Db	408	TATTGGAGACAAATCGTAAATATGTCTTTTGGAGTGTCTCAGTGCACGAAGATGTTCCG	467
Qy	147	SerSerArgLeuIleArgGluGluGluMetGluAspAlaIleThrPheLeuArgSerLys	166
Db	468	ACATTTAGCTCTATTAGCGGAATGAAGTCTCTCGTCTCATTAATTTATCCGGTCACT	527
Qy	167	AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleSerIleMetIle	186
Db	528	TCTGGTGAACTATTAAATGTTACGAAAGGATCTTTTGTTCACAAGCTCCATGACATGT	587
Qy	187	ArgThrSerValGlyAsnCys--LysGlnLysGluArgLeuLeuSerValAlaAspAla	205
Db	588	AGATCAGCGTTGGCGCAAGTGTTCANAAGACACGAACAAATATTATCAACTAATAAAGAA	647
Qy	206	ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu	225
Db	648	GTGATACTTTAGCAGAGGTTTGATGTGGCTGCATATTTCCCTTCACATGAAGTTCTTT	707
Qy	226	HisTyrIleIleGlyAlaGluSerLysProArgLeuHisGlnGluIleAspAspIle	245
Db	708	CATGTCTCAGTGGAAATGAAGGTAAGATTATGAATGCACACCAATAAGGTAGATGCCATT	767
Qy	246	LeuGluGluIleLeuAsnGluHis-----LysAlaAsnLysPro	258
Db	768	TTTGAGAATGTCATCAATGACACACGAACAAAATCTTGCAATTTGGAAAAAATTAATGGAGCG	827
Qy	259	PheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValPro	278
Db	828	TTAGAGGTGAAGATTAAATTGATGTTCTCTAAGACTTATGAATGATGAGGCCCTTCAA	887
Qy	279	ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer	298
Db	888	TTTCTCATCACAACGACAACATCAAGGCCATAATTTTGACATGTTTGCTGCGCGGACA	947
Qy	299	GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu	318
Db	948	GAGACTTCATCGTCAACAATTTGTGTGGCTATGCTAGAAAATGGTGTGAAAAATCCACCGCGTA	1007
Qy	319	LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu	338
Db	1008	TTCCGCAAAAGCTCAAGCAGAAGTAGAAGAAGCATTTTAGAGAAAAGAAACTTTTCGATGAA	1067
Qy	339	SerArgPheHisAspLeuLysPhePheLysLeuValLysGluThrLeuArgLeuHis	358
Db	1068	AATGATGTGGAGGAGCTAACTACTATAAGTTAGTAAATAAAGAAACTCTAAGACTTCAT	1127
Qy	359	ProProVal---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyr	377
Db	1128	CCACCGGTTCCACTTTTGCTCCCAAGAGAATGTAGGGAAGAGACAAATAATAAACCGCTAC	1187
Qy	378	GluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsn	397
Db	1188	ACTATTCTGTAAAGACCAAAAGTCATGGTTAAATGTTTGGCGCTTTGGGAAGAGATCCAAAA	1247
Qy	398	ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr	417
Db	1248	TATTGGAATGACGCAGAAACTTTTATGCCAGAGAGATTGAGCAGTGTCTTAGAGATTTT	1307
Qy	418	LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIle	437
Db	1308	GTTCGTATAATATTTGAATATCTTCCATTTGCTGGCGAAGAGAGATTTGCTCTGGGATT	1367

Qy	438	ThrSerAlaIleThrAenLeuGluTyrValIleAenLeuLeuTyrHisPheAsnTyr	457
Db	1368	TCGTTTGGCTTAGCTAATGCTTATTTGGCAATGGCTCAATTACTATATCATCTACTTCGATTGG	1427
Qy	458	GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAla	477
Db	1428	AAACTCCCTGCTGGAATCGAACCAAGCGACTTGGACTTGGATGGTTGGTTGGAGTAAC	1487
Qy	478	LeuArgTyrLysIleAspLeuLysLeuIleProIleProTyrGln	492
Db	1488	GCCCTAGAAAAGTGCCTTACTTGGTTGGCGACTCTTATCAA	1532
RESULT 6			
LOCUS	CS223962	Sequence 218 from Patent WO2005111217.	1673 bp DNA linear PAT 15-DEC-2005
DEFINITION	CS223962		
ACCESSION	CS223962		
VERSION	CS223962.1	GI:83685860	
KEYWORDS			
SOURCE			
ORGANISM		Nicotiana tabacum (common tobacco)	
REFERENCE			
AUTHORS		Xu, D.	
TITLE		Nicotiana nucleic acid molecules and uses thereof	
JOURNAL		Patent: WO 2005111217-A 218 24-NOV-2005;	
FEATURES		U.S. Smokeless Tobacco Company (US)	
source		Location/Qualifiers	
		1..1673	
		/organism="Nicotiana tabacum"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:4097"	
ORIGIN			
Alignment Scores:			
Pred. No.:	7.44e-122	Length:	1673
Score:	1148.00	Matches:	222
Percent Similarity:	63.2%	Conservative:	91
Best Local Similarity:	44.8%	Mismatches:	166
Query Match:	44.6%	Indels:	16
DB:	2	Gaps:	5
US-10-759-813-2 (1-500) x CS223962 (1-1673)			
Qy	12	LeuIleSerPheLeuLeuValLeuIleLeuValValMetArgLeuTyrLysGln	31
Db	54	CTGGTTTCCATTTCCTATTCTCTTCTCTTTTGGTAAGGTATGGAAGAACTCC	113
Qy	32	AsnPro-----ProProGlyProTyrLysPheProIleIleGlyAsnLeu	46
Db	114	AATAGCCAAAGCAAAAAGTTGCCACCAGGTCCATGGAACATACCAATACTAGGAAGTATG	173
Qy	47	ProHisLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle	66
Db	174	CTTCATATGGTT-----GGTGACTACCAACCATGTCTCTTAGAGATTAGCCAAAAA	227
Qy	67	TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValIleSerSerAla	86
Db	228	TATGGACCACTTATGCACCTTCAATTAGGTGAAGTTCTCGCGTGTGGTTACTTCTCCT	287
Qy	87	GluAlaIalysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal	106
Db	288	GATACGGCAAAAGAAGPATTTAAAAAATCATGCATCGCTTTTCGGTCTAGGCCTTAGCCCT	347
Qy	107	LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp	126
Db	348	TTGGCCCGGAGATTGTCTGTACATAGTCTGAICTAGCCCTTTTGGCCCTATGGCGAC	407
Qy	127	HistTArgGlnMetLysIleTyrPileLeuGluPheLeuSerAlaLysValGln	146

```

Db 408 TATTGGAGACAAATCGCTAAATATATGCTTTGGAAAGTGTCTCAGTGCCCAAGAAATGTCGG 467
QY 147 SerSerArgLeuLeuArgGluGluMetGluuAspAlaIleThrPheLeuArgSerLys 166
Db 468 ACATTAGCTCTATTAGCGGAATGAAGTCTTCGTCTCATTAATTTTATCCGGTCATCT 527
QY 167 AlaGlySerProValAsnIleThrLysIleIleThrGlyIleIleSerIleMetIle 186
Db 528 TCTGGTGAACCTATTATGTTACGGAAGGATCTTTTGTTCACAAGCTCCATGACATGT 587
QY 187 ArgThrSerValGlyAsnCys---LysGlnLysGluArgLeuLeuSerValAlaAspAla 205
Db 588 AGATACGCTTTGGCAAGTGTCAAAGAGCAAGCAAAATTTATACAATTAATAAGAA 647
QY 206 ValAsnGluAlaIleThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu 225
Db 648 GTGATACTCTTAGCAGGAGGTTTGATGTGGTGACATATTCCTTCATGAAATTCTT 707
QY 226 HisTyrIleGlyAlaGluSerLysProArgArgLeuHisGlnGluIleAspIle 245
Db 708 CATGTGCTCAGTGAATGAAGGTAAAGATTATGAATGCACACCATAAGGTAGATGCCATT 767
QY 246 LeuGluGluIleLeuAsnGluHis-----LysAlaAsnLysPro 258
Db 768 GTTGAAGATGTCATCAATCAGCACAGCAAGAAAAATCTTTGCAATTTGGGAAAACTAATGGAGCG 827
QY 259 PheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValPro 278
Db 828 TTAGGAGGTGAAGATTATTAATGATGTTCTTAGACTTATGAATGATGAGCGCTTCAA 887
QY 279 ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer 298
Db 888 TTTTCTTATCACCACCAACATCAAGCTATAATTTTGTGACATGTTTCTGCTCGCGGACA 947
QY 299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
Db 948 GAGACTTCATCGTCAACAATTTGTGTGGCTATGGTAGAATGGTGAATAATGCTCAGCGGTA 1007
QY 319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
Db 1008 TTCGGGAAGCTTCAGCAGAGTAGAGGAAGCATTTAGAGGAAAGAAACTTTTCGATGAA 1067
QY 339 SerArgPheHisAspLeuLysPheLysLeuValValLysGluThrLeuArgLeuHis 358
Db 1068 AATGATGTGAGGAGCTAACTAACTAAAGTTAGTAATAAAGAAACTCTAAGACTTCAT 1127
QY 359 ProProVal---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyr 377
Db 1128 CCACCGGTTCACATTTTGTCTCCCAAGAGAAATGTAGGGAAGAGACAAATATAACCGCTAC 1187
QY 378 GluIleHisProAsnThrArgIleValValAsnAlaIleAlaIleGlyArgAspProAsn 397
Db 1188 ACTATTCTGTAAAGCAAAAGTCATGTTATATGTTGGCTTTGGGAAGAGATCCAAA 1247
QY 398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr 417
Db 1248 TATTGGAATGACGCAAACTTTTATGCCAGAGAGATTTGAGCAGTGCTCTAAGGATTTT 1307
QY 418 LysGlyThrThrPheGluLeuValProPheGlyValaGlyLysArgIleCysProGlyIle 437
Db 1308 GTTGGTAATAATTTTGAATATCTCCATTTGGTGGCGGAAGGAGGATTTGTCCTGGGATT 1367
QY 438 ThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyHisPheAsnTrp 457
Db 1368 TCGTTTGGCTTAGCTTAATGCTTATTGGCATTTGGCTCAATTAATATATCACTTCGATGG 1427
QY 458 GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAla 477
Db 1428 AAATCTCCCTGCTGGAATCAACCAAGCGACTTTGGACTTGCATGAGTTGGTGGTAAGT 1487
QY 478 LeuArgLysLysAlaLeuAspLeuLysLeuIleProIleProTyrGln 492
Db 1488 GCCGCTAGAAAAAGTGACCTTTTACTTTGGTTGCGACTCTCTTATCA 1532

```

RESULT 7

CS223948
LOCUS CS223948 1657 bp DNA linear PAT 15-DEC-2005
DEFINITION Sequence 204 from Patent WO2005111217.
ACCESSION CS223948
VERSION CS223948.1 GI:83685853
KEYWORDS
SOURCE Nicotiana tabacum (common tobacco)
ORGANISM Nicotiana tabacum

REFERENCE

1 Xu, D.
AUTHORS
TITLE Nicotiana nucleic acid molecules and uses thereof
JOURNAL Patent: WO 2005111217-A 204 24-NOV-2005;
U.S. Smokeless Tobacco Company (US)
FEATURES
Location/Qualifiers
1. .1657
/organism="Nicotiana tabacum"
/mol_type="unassigned DNA"
/db_xref="taxon:4097"

ORIGIN

Alignment Scores:

Pred. No.: 2,13e-121 Length: 1657
Score: 1144.00 Matches: 223
Percent Similarity: 66.5% Conservative: 107
Best Local Similarity: 45.0% Mismatches: 152
Query Match: 44.4% Indels: 14
DB: 2 Gaps: 6

US-10-759-813-2 (1-500) x CS223948 (1-1657)

```

QY 10 SerIleLeuIleSerPheLeuLeuValLeu---IleLeuValValValMetArgLeuTrp 28
Db 37 TCTAACTTAGTTGCAATCTTCTCTCTCTCAGCATCTTTCTTCTTATTCAAAAATGG 96
QY 29 Lys-----LysGlnAsnProProGlyProTrpLysPheProIleIleGlyAsnLeu 46
Db 97 AAAACCAAGAAACTAAATTTGCTCTCTGTCATGAAATTTACCTTTTATGGAAAGTTA 156
QY 47 ProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle 66
Db 157 CACCATTTGGCTGTGGCAGGTCCACTTCTCTACCATGGCTTAAAAAATTTAGCCAAACGC 216
QY 67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerAla 86
Db 217 TATGTCCTCTTATGCATTTACAACTTGGACAAATTTCTTACACTCGTCATCATCACCT 276
QY 87 GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
Db 277 CAATGGCAAAAGAGTAGTACTAAAAAATCTCAGACCTCGCTTTTGGCCACTAGACCAAGCTT 336
QY 107 LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
Db 337 GTCTGTGGCGACATCATTCACCTACGACGACGACATAGCATTTTCGCCATACGGTGA 396
QY 127 HisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln 146
Db 397 TACTGGAGACAAATTCGTAATAATTTGCATATTGCAACTCTTGAGTGCCCAAGATGGTCAAG 456
QY 147 SerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLys 166
Db 457 TTTTGTAGTCGATTCGCCAAGATGAGCTCTCGAAGATGGTTTCATCTATACGACGACG 516
QY 167 AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleIleSerIleMetIle 186
Db 517 CCCATCTTCAGTCAATCTTACCGACAAGATTTTGTGTTTACGAGTTTCGGTAATTTGT 576
QY 187 ArgThrSerValGlyAsn---CysLysGlnLysGluArgLeuLeuSerValAlaAspAla 205

```



```

Db      768 GTTGAGAAATGTCATCAATGAGCACAAAGAAAAATCTTGCATTTGGGAAAACTAATGAGCG 827
Qy      259 PheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValPro 278
Db      828 TTAGGAGGTGAAGATTATTAATGATGTTCTCTAAGACTTATGATGATGAGGCCCTTCAA 887
Qy      279 ValProValThrAsnGluSerLeuAlaSerValLeuGlnMetPheThrAlaGlySer 298
Db      888 TTTCTCTATCACCACCAACATCAAGCTATAAATTTTGGACATGTTTGGCTGCGCGGACG 947
Qy      299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
Db      948 GAGACTTCACTGCAACAATTTGTGGCTATGTTAGTAATAATGGTGAATAATCCAGCCGTA 1007
Qy      319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
Db      1008 TTCGCGAAGCTCNAGCAGAGTAAAGAGCATTTAGAGGAAAAGAACTTTCGATGAA 1067
Qy      339 SerArgPheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHis 358
Db      1068 AATGATGTGGAGGAGCTAACTACCTAAAGTTAGTAATAAAGAACTCTAAGACTTCAT 1127
Qy      359 ProProVal--ValleulleProArgGluCysArgGluThrThrArgileAspGlyTyr 377
Db      1128 CCACCGGTTCACCTTTTGTCTCCCAAGAGATGTAGGGAAGAGACAAATATAAACGGCTAC 1187
Qy      378 GluileHisProAsnThrArgileValValAsnAlaTipAlaileGlyArgAspProAsn 397
Db      1188 ACTATTCCTGTAAGCCAAAGTCATGTTAATGTTGGCTTTGGGAAGAGATCCAAA 1247
Qy      398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaileAspTyr 417
Db      1248 TATTGGAATGACGAGAACTTTTATGCCAGAGAGATTTGAGCAGTGCTCTAAGGATTT 1307
Qy      418 LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgileCysProGlyIle 437
Db      1308 GTTGGTAATAATTTTGAATATCTTCATTTGTGGCGGAAGAGGATTTGTCTGGGATT 1367
Qy      438 ThrSerAlaileThrAsnLeuGluTyrValIleleAsnLeuLeuTyrHisPheAsnTyr 457
Db      1368 TCGTTTGGCTTAGCTAATGCTTATTGTCATTTGGCTCAATTAATATATCACTTCGATTGG 1427
Qy      458 GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaileGlyGlyAla 477
Db      1428 AAACCTCCCTGCTGGAATCAACCAAGCAGCTTGGACTGACTGAGTTGGTGGAGTAAC 1487
Qy      478 LeuArgLysLysIleAspLeuLysLeuIleProIleProTyrGln 492
Db      1488 GCCGCTAGAAAAAGTACCTTTTACTTGGTTGCGACTCCTTATCAA 1532

RESULT 9
STU296346
LOCUS Solanum tuberosum mRNA for cytochrome P450 (CYP71D4 gene).
DEFINITION
ACCESSION AJ296346
VERSION AJ296346.1 GI:12331297
KEYWORDS CYP71D4 gene; cytochrome P450.
SOURCE Solanum tuberosum
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum.

REFERENCE
1
AUTHORS Schneider M.
TITLE Cloning of a potato gene (CYP71D4) induced after infection with
Phytophthora infestans
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 1701)
AUTHORS Schneider M.
TITLE Direct Submission
JOURNAL Submitted (18-JAN-2001) Schneider M., Plant Biology, University of
Fribourg, 3, rue Albert Gockel, CH-1700 Fribourg, SWITZERLAND
LOCATION/Qualifiers

```

```

1. .1701
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Datura"
/db_xref="taxon:4113"

```

Gene

```

1. .1701
/gene="CYP71D4"
23. .1531
/gene="CYP71D4"
/function="unknown"
/codon_start=1
/product="cytochrome P450"
/protein_id="CAC24711.1"

```

CDS

```

/db_xref="GI:12331298"
/db_xref="GOA:O9AVQ2"
/db_xref="InterPro:IPR001128"
/db_xref="InterPro:IPR002401"
/db_xref="UniProtKB/TrEMBL:O9AVQ2"
/translation="MQFUSLASIFLFLRKKWNSNSQSKLPPGPKWKLPLGSG
MLHMAGLPHVLRDLAKKYGGLHMLQLGEVAVVVTSPDMAKEVLKTHDIAFARPK
LLAPEIVCVNRSDIAFCPYGDVYRWQRKICVLELLSAKVRSYGSIRRDVGLVNFPI
RSSGPEVNFTEERLEFLFTSSMTCSRAGKVPKEODKFIOLIKEVIGLAGFDVADLFP
SLKFLHVLGSMGKIMNAHKKVDAIVEDVINEHKKKFAIKGTNGALGGDLIDLVLRL
MNDGLQFPITNDNIKAIIFDMAAGTETSSSTLVAMVQMMKPSVIAKAAQAEVREA
FKDKETFDENVDEELKLVIKETLRHPVPLLVPRECREETDINGTIPVKTKVM
VNWALGRDPKYWDDAESPKPERFQORSVDVGNPNPEYLPFGGRRICPGISPLNAV
YLPLAHLHYHFDWKLPIGMEPKDLNLTELGVTVAAKDDLLVLATPYPPRQ"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 3,76e-121 Length: 1701
Score: 1142.00 Matches: 217
Percent Similarity: 64.2% Conservative: 101
Best Local Similarity: 43.8% Mismatches: 161
Query Match: 44.4% Indels: 16
DB: 4 Gaps: 5

```

US-10-759-813-2 (1-500) x STU296346 (1-1701)

```

Qy      12 LeuileSerPheLeuLeuValLeuileLeuValValMetArgLeuTrpLysGln 31
Db      38 TTGGCTTCCATCTTCTTTCTATCTTTTCTTAAAGAAATGGAAAACTCG 97
Qy      32 AsnPro-----ProGlyProTrpLysPheProIleileGlyAsnLeu 46
Db      98 ATAGCCCAATCGAAAAAATGGCTCCAGTCCATGGAACTTCCTTACTAGGAAGATG 157
Qy      47 ProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle 66
Db      158 CTTTCAT-----ATGCTGGTGGACTTCCACACCATGCTCTTAGAGATTAGCCAAAAA 211
Qy      67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSerAla 86
Db      212 TATGACCACCTTATGATCTTCAACTGGTGAAGTCTCTGCAGTTGTAGTAACCTTCCT 271
Qy      87 GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
Db      272 GATATGGCGAAGAAGTACTAAAACTCATGACATCGCTTTTCGCTCTAGGCTAACTT 331
Qy      107 LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
Db      332 TTGGCCCGGAAATTTGTTTGTACCAACAGGTCTGACATTTGCGCTTTTGGCCCTACGAGAT 391
Qy      127 HisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln 146
Db      392 TACTGGAGACAAATCGTAAAAATTTGTGTCCTTGGAAATTTGTTGAGTCCCAAGATGTC 451
Qy      147 SerSerArgLeuileArgGluGluMetGluAspAlaIleThrPheLeuArgSerLys 166
Db      452 TCATATGCTCGATTAGCGCGCATGAGTTGATCGCTTGTAAATTTATTCGGTCATCT 511
Qy      167 AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleIleSerIleMetIle 186

```

```
Db 512 TCGGGTGAGCCGGTTAAATTTACTGAAAGGTTGTTTTTTTTCACAAAGTTCAATGACATGT 571
Qy 187 ArgThrSerValGlyAsnCys---LysGlnLysGluArgLeuLeuSerValalaAspAla 205
Db 572 AGATCAGCGTTCGGGAAGAGTGTCAAGAACAGGACAAATTTATACAACTAATCAAGAA 631
Qy 206 ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu 225
Db 632 GTGATTGGTTAGCAGGAGGATTTGATGGCTGATATCTTCCCATCAATTTGAAGTTTCTC 691
Qy 226 HisTyrIleIleGlyAlaGluSerLysProArgLeuHisGlnGluIleAspAspIle 245
Db 692 CATGTCCTTAGTGGATGAAGGTAAATTTATGAACGCTCATCATAGTAGTACCAATT 751
Qy 246 LeuGluGluIleLeuAsnGluHis-----LysAlaAsnLysPro 258
Db 752 GTTGAAGATGTCATCAATGAGCACAAAGAAATTTGCAATTTGGGAAACAAATTAATGGTGA 811
Qy 259 PheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValPro 278
Db 812 TTAGGTGGTGAAGATCTAATTCATGCTCCTTAAAGACTTTATGAATGATGGAGGCTTCAA 871
Qy 279 ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer 298
Db 872 TTTCCGATCACCACCAACATCAAGCTATTATTTTCGACATGTTTGTGTCAGAAACA 931
Qy 299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
Db 932 GAGACTTCATGCTCAACACTTGTCTGGCAATGGTGCAAAATGATGAAACCCAAAGTGTA 991
Qy 319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
Db 992 ATCGCAAGCTCAAGCAGAGTGGGAGAGCCTTTAAAGACAAAGAAACCGTTTCGATGAA 1051
Qy 339 SerArgPheHisAspLeuLysPheLysValLeuValLysGluThrLeuArgLeuHis 358
Db 1052 AATGATGTAGAGGAGCTGAATATCTTAAAGTTAGTCATTAAAGAAACTCTAAGACTCCAT 1111
Qy 359 ProProVal---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyr 377
Db 1112 CCACAGTTCACACTTTTGTCCCAAGAGATGTAGGGAAGACGGATATAAACGGCTAC 1171
Qy 378 GluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsn 397
Db 1172 ACTATTCTCTGTGAAGACCAAAAGTCATGTTATGTTTGGGCATTTGGGAAGAGATCCGAAA 1231
Qy 398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr 417
Db 1232 TATTGGGATGATGCAGAAAGTTTAAAGCCAGAGAGATTTGACGCGCTCTGTGCACCTT 1291
Qy 418 LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIle 437
Db 1292 GTTGTAACAAATTTTGAGTATCTTCCCTTTGGCGGTGGGAGAAGATTGTCCCGGATA 1351
Qy 438 ThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrp 457
Db 1352 TCATTGGCTTAGCTAAATGTTATTGTCGGTTGGCTCATTTGTTATATCATCTTCACCTGG 1411
Qy 458 GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyAla 477
Db 1412 AAATCTCCCTATTGGAAATGAGCCAAAGACTTGAACCTTGACTGAATTTGGTTGGAGTAAC 1471
Qy 478 LeuArgLysIleAspLeuLysLeuIleProIleProTyrGln 492
Db 1472 GCTGCCAGAAAAGATGACCTTATTTTGGTTGCCACTCTCTATGAA 1516
```

RESULT 10

```
CS223950
LOCUS          CS223950          1610 bp          DNA          linear          PAT 15-DEC-2005
DEFINITION    Sequence 206 from Patent WO200511217.
ACCESSION    CS223950
VERSION      CS223950.1  GI:83685954
KEYWORDS
```

```
SOURCE      Nicotiana tabacum (common tobacco)
ORGANISM    Nicotiana tabacum
REFERENCE   1
AUTHORS     Xu,D.
TITLE       Nicotiana nucleic acid molecules and uses thereof
JOURNAL     Patent: WO 200511217-A 206 24-NOV-2005;
            U.S. Smokeless Tobacco Company (US)
FEATURES    Location/Qualifiers
             source          1..1610
                        /organism="Nicotiana tabacum"
                        /mol_type="unassigned DNA"
                        /db_xref="taxon:4097"
ORIGIN
Alignment Scores:
Pred. No.:      6.51e-120      Length:      1610
Score:          1131.00      Matches:    222
Percent Similarity: 66.1%      Conservative: 105
Best Local Similarity: 44.8%      Mismatches: 154
Query Match:    43.9%      Indels:      14
DB:             2           Gaps:        6
US-10-759-813-2 (1-500) x CS223950 (1-1610)
```

```
Qy 10 SerIleLeuIleSerPheLeuLeuValLeu---IleLeuValValMetArgLeuTrp 28
Db 37 TCTAACTTAGTTCGATCTTCTCTCTCTCCAGCATCTTTCTTGTTATTCAAAAATGG 96
Qy 29 Lys-----LysGlnAsnProProGlyProThrLysPheProIleIleGlyAsnLeu 46
Db 97 AAAACCCAGAAATAAATTTCCTCTCTGTCATGGAATAATTACCTTTATTTCGAAAGTTTA 156
Qy 47 ProHisIleLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle 66
Db 157 CACCATTTGGCTGTGGCAGGTCCACTTCTCCATTCGCCTTAAAAAATTTAGCCAAACGC 216
Qy 67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSerAla 86
Db 217 TATGGTCTCTTATGTCATTTACAACTTTCGACAAATTCCTACACTCGTCATATCATCACCT 276
Qy 87 GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
Db 277 CAATATGGCAAAAGAGTACTAAATACTCAGCACTCGCTTTTGGCCTAGACCAAGCTT 336
Qy 107 LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
Db 337 GTCGTGGCCGACATCATTCACACAGCAGCAGACATAGCACCTTTCGCCATACCGTGAA 396
Qy 127 HisTrpArgGlnMetLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln 146
Db 397 TACTGGAGACAAATTCGTAAAAATTTGCATATTGGAACCTTTGAGTGCCCAAGATGTCGAA 456
Qy 147 SerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLys 166
Db 457 TTTTATTAGCTCGATTCGCCAAGATGAGCTCGAAGATGGTTTTCATATACGAAACGAG 516
Qy 167 AlaGlySerProValAsnIleThrLysIleTyrGlyIleIleIleSerIleMetIle 186
Db 517 CCCAATCTTCAGTCATCTTACCAGCAAGATTTTTCGTTTACGAGTTCGGTAATTTGT 576
Qy 187 ArgThrSerValGlyAsn---CysLysGlnLysGluArgLeuLeuSerValalaAspAla 205
Db 577 AGATCAGCTTTTAGGGAAGATATGTGTCACCAAGACAAATTCATTTTTTATGAGGGAA 636
Qy 206 ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu 225
Db 637 ATAATATCATTTGGCAGGTGGATTTAGTATTGCTGATTTTTCCTACATGGAAATGATT 696
Qy 226 HisTyrIleIleGlyAlaGluSerLysProArgLeuHisGlnGluIleAspAspIle 245
```



```

Db      1061 CAGATTGATCTTGATGAGCTAAATATCTCAAGTTAGTAAATCAAGAAACCTTTAAGAATG 1120
Qy      HisProProVal--ValLeuLeuProArgGluCysArgGluThrThrArgLeuAspGly 376
        |||||
Db      1121 CACCCTCCAAATCTCTATAGTTCTTAGAGAATGATGGAGATCAAGAATGATGATGGT 1180
Qy      TyrGluLeuHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspPro 396
        |||||
Db      1181 TACAATATACCTTTCAAAACACAGAGTCATAGTTAATGTCATGGCAATCGGACGAGATCCA 1240
Qy      AsnThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAsp 416
        |||||
Db      1241 GAAAGTTGGAGATACCCCGAAAGCTTTATGCCAGAGAGATTTGAGAGATAGTTCTATTGAC 1300
Qy      TyrLysGlyThrThrPheGluLeuValProPheGlyValGlyLysArgIleCysProGly 436
        |||||
Db      1301 TTCTTTGGAAATCATCATGTTTATACCATTTTGGTGAGAGAGAGATTTGTCGGGA 1360
Qy      IleThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsn 456
        |||||
Db      1361 ATGCTATTGGTTTAGCTAATGTTGACCAACCTTTAGCTCAGTTACTTTATCATCTTCGAT 1420
Qy      TrpGluLeuAlaAspGlyLeuThrProGlnThrLeuAspMetThrGluAlaIleGlyGly 476
        |||||
Db      1421 TGGAAATCCCTAATCGCAAAAGTCATGAGATTTTCGACATGACTGAGTCACCTGGAATT 1480
Qy      AlaLeuArgLysLysIleAspLeuLysLeuIleProIleProTyr 491
        |||||
Db      1481 TCTGCTACAGAAAGGATGATCTGTTTGTATGTCACCTCTTAT 1525

```

RESULT 13

```

CS223944 LOCUS 1566 bp DNA linear PAT 15-DEC-2005
DEFINITION Sequence 200 from Patent WO2005111217.
ACCESSION CS223944
VERSION CS223944.1 GI:83685851
KEYWORDS
SOURCE Nicotiana tabacum (common tobacco)
ORGANISM Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Nicotiana.
1
AUTHORS Xu,D.
TITILE Nicotiana nucleic acid molecules and uses thereof
JOURNAL Patent: WO 2005111217-A 200 24-NOV-2005;
U.S. Smokeless Tobacco Company (US)
FEATURES
source
1. .1566
Location/Qualifiers
/mol_type="unassigned DNA"
/db_xref="taxon:4097"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 1,22e-119 Length: 1566
Score: 1128.50 Matches: 218
Percent Similarity: 66.1% Conservative: 109
Best Local Similarity: 44.0% Mismatches: 155
Query Match: 43.8% Indels: 13
DB: 2 Gaps: 5

US-10-759-813-2 (1-500) x CS223944 (1-1566)

Qy      10 SerIleLeuIleSerPheLeuValLeu---IleLeuValValValMetArgLeuTrp 28
        |||||
Db      29 TCTAATTAGTTCGATCTTGTCTTCTCTCCAGCATCTTCTTCTTATTTCAAAAATGG 88
Qy      29 Lys-----LysGlnAsnProProGlyProTrpLysPheProIleIleGlyAsnLeu 46
        |||||
Db      89 AAACACGAAACAACTAAATTTGCTCTGCTCATGGAATTAACCTTTTATGGAAGTTTA 148
Qy      47 ProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle 66

```

```

Db      149 CACCAATTTGGCTGTGCAGGTCACCTTCTCCACCATGGCTAAATAATTTAGCAACGC 208
Qy      TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSerAla 86
        |||||
Db      209 TATGGTCTCTTATGATTTTACAACCTTGGACAAATTCCTACACTCATCATCATCACCT 268
Qy      GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
        |||||
Db      269 CAAATGCGAAAGAAGTACTTAAAACTCAGCACCTCGCTTTTCCCACTAGCAACAGCTT 328
Qy      LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
        |||||
Db      329 GTCGTGGCCGACATCATCTACGACAGCAGCAGCATAGCATTTTCTCCGTACGGTAA 388
Qy      HisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysValGln 146
        |||||
Db      389 TACTGGAGACAAATTCGTAAATTTGCATATTTGGAACCTCTTGAGTCCCAAGATGGTCAA 448
Qy      SerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLys 166
        |||||
Db      449 TTTTGTAGTCTGATTCGCAAGATGAGCTCTCGAAGATGCTCTCATCTATACGAACGACA 508
Qy      AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleIleSerIleMetIle 186
        |||||
Db      509 CCAATCTTTACAGTCAATCTTACTGACAAAATTTTGTGTTTACGAGTTTCGGTAACCTGT 568
Qy      ArgThrSerValGlyAsn---CysLysGlnLysGluArgLeuLeuSerValAlaAspAla 205
        |||||
Db      569 AGATCAGCTTTTAGGGAAGATATGTGTGACCAAGACAAATTCATATTTTATGAGGAA 628
Qy      ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu 225
        |||||
Db      629 ATAAATATCATTCGCGAGTGGATTTAGTATTGCTGATTTTTCCTACATGGAAATGATT 688
Qy      HisTyrIleIleGlyAlaGluSerLysProArgArgLeuHisGlnGlnIleAspAspIle 245
        |||||
Db      689 CATGATATTGATGGTTTGGAAATCTAAACTGGTGAAAGCACATCGTAAGATTTGATGAA 748
Qy      LeuGluGluIleLeuAsnGluHisLysAlaAsnLys----- 257
        |||||
Db      749 TTGGGAAATTTGTTGTATGAGCACAAAAGACAGACAGATGGCAAGAGGTAATGGT 808
Qy      ProPheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnVal 277
        |||||
Db      809 GAATTTGGTGGTGAAGATTGATGATGATTTGTTTAAAGAGTTAGAGAAAGTGGAGAGTT 868
Qy      ProValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGly 297
        |||||
Db      869 CAAATTCCTATCACAAATGACAATATCAATCAATTAATTAATCGACATGTTCTCTCGCGGA 928
Qy      SerGluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThr 317
        |||||
Db      929 TCTGAACATCATCGACGACTATAATTTGGGCAATTAGCTGAAATGATGAAGAACCAAGT 988
Qy      GluLeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAsp 337
        |||||
Db      989 GTTTTAGCAAGCCCAAGCTGAAGTAAGCGAAGCTTTTGAAGAGAGAAAAGGTTTTCAA 1048
Qy      GluSerArgPheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeu 357
        |||||
Db      1049 CAGATTGATCTTGATGAGCTAAATATCTCAAGTTAGTAAATCAAGAAACCTTTAAGAATG 1108
Qy      HisProProVal---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGly 376
        |||||
Db      1109 CACCCTCCAAATTCCTTATAGTTTCTAGAGAATGATGAGAGGATACAAAGATTTGATGT 1168
Qy      TyrGluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspPro 396
        |||||
Db      1169 TACAATATACCTTTCAAAACAAAGATCATAGTTAATGATGGCAATCGACAGATCCA 1228
Qy      AsnThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAsp 416
        |||||

```

```

Db      1229  GAAAGTTGGGATGACCCCGAAAGCTTTATGCGCAGAGAGATTGGAATAGTTCATTGAC 1289
Qy      417  TyrLysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgileCysProGly 436
Db      1289  TTTCTTGGAATCATCATCAGTTATACCAATTTGGTGCGAGGAAGGATTTGTCGGGA 1348
Qy      437  IleThrSerAlaIleThrAsnLeuGluTyrValIlelleAsnLeuLeuTyrHisPheAsn 456
Db      1349  ATGCTATTGTTGGTTAGCTAATGTCGACAAACCTTTTAGCTCGACTTACTTATCATCTCGAT 1408
Qy      457  TrpGluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGly 476
Db      1409  TGGAACTCCCTAATGGCAAGTCAATGAGAAATTCGACATACCTGAGTCACTCGGAATT 1469
Qy      477  AlaLeuArgLysIleAspLysLeuIleProIleProTyr 491
Db      1469  TCTGCTACAGAAGAGGATGATCTTGTGTTGATTCGCCTCTTAT 1513

RESULT 14
CS223904      1581 bp      DNA      linear      PAT 15-DEC-2005
LOCUS      CS223904
DEFINITION      Sequence 160 from Patent WO2005111217.
ACCESSION      CS223904
VERSION      CS223904.1  GI:83685433
KEYWORDS
SOURCE
ORGANISM
Nicotiana tabacum (common tobacco)
Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiales; Solanales; Solanaceae; Nicotiana.
1
REFERENCE
AUTHORS      Xu, D.
TITLE      Nicotiana nucleic acid molecules and uses thereof
JOURNAL      Patent: WO 200511217-A 160 24-NOV-2005;
U.S. Smokeless Tobacco Company (US)
FEATURES
source
1..1581
/organism="Nicotiana tabacum"
/mol_type="unassigned DNA"
/db_xref="taxon:4097"
ORIGIN

Alignment Scores:
Pred. No.:      1.23e-119      Length:      1581
Score:      1128.50      Matches:      218
Percent Similarity:      66.1%      Conservative:      109
Best Local Similarity:      44.0%      Mismatches:      155
Query Match:      43.8%      Indels:      13
DB:      2      Gaps:      5

US-10-759-813-2 (1-500) x CS223904 (1-1581)
Qy      10  SerIleLeuIleSerPheLeuLeuValLeu---IleLeuValValValMetArgLeuTrp 28
Db      46  TCTAACTTAGTTCGATCTTCTGCTCTCTCTCTCCAGCATCTTCTCTATTCAAAAATGG 105
Qy      29  Lys-----LysGlnAsnProProGlyProTrpLysPheProIleIleGlyAsnLeu 46
Db      106  AAAACCAGAAAACTAAATTTGCTCTCTGCTCCATGGAAATTAACCTTTTATTGGAAAGTTA 165
Qy      47  ProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuIleGlnIle 66
Db      166  CACCAATTGGCTGTGGCAGGTCCACTTCTCCACCATGGCCTTAAAAAATTTAGCCAAACGC 225
Qy      67  TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValIleSerSerAla 86
Db      226  TATGTCCTCTTATGCAATTACACTTGGACAAATTCCTACACTCATCATCATCACT 285
Qy      87  GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
Db      286  CAAATGGCAAAAGAGTACTAAAAACTACGACCTCGCTTTTGGCCACTAGACCAAAGCTT 345
Qy      107  LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126

```

```

Db      346  GTCTGCCCGCATCATCTCACTACGACAGCAGCATAGCATTTTCTCCGTACCGTGAA 405
Qy      127  HisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln 146
Db      406  TACTGGAGACAAATTCGTAAATTTGCATATTGGAATCTTTGAGTCCCAAGATGTCAAA 465
Qy      147  SerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLys 166
Db      466  TTTTGTAGTCGATTCGCAAGATGAGCTCTCGAAGATGCTCTCATCTATATACGACGACA 525
Qy      167  AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleIleSerIleMetIle 186
Db      526  CCCAATCTTACAGTCAATCTTACGACAAAATTTTGTGTTTACAGTTCGGTAACTTGT 585
Qy      187  ArgThrSerValGlyAsn---CysLysGlnLysGluArgLeuLeuSerValAlaAspAla 205
Db      586  AGATCAGCTTTAGGGAAGATATGTTGGTACCAACAAATTTGATCATTTTATGAGGGA 645
Qy      206  ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu 225
Db      646  ATAATATCATTTGGCAGGTGGATTTAGTATTGCTGATTTTCCCTACATGGAAATGATT 705
Qy      226  HisTyrIleIleGlyAlaGluSerLysProArgLeuHisGlnGluIleAspAspIle 245
Db      706  CATGATATTGATGTTGGAATCTAAACTGTGTGAAAGCACATCGTAAGATTTGATGAATT 765
Qy      246  LeuGluGluIleLeuAsnGluHisLysAlaAsnLys----- 257
Db      766  TTGGAAATGTTGTTGATGACACAAAAGAACAGACAGCATGGCAAGAGGTAATGTT 825
Qy      258  ProPheGluAlaAspAsnLeuMetAspValLeuLeuLeuGlnLysAsnGlyAsnVal 277
Db      826  GAAATTTGGTGTGAAGATTTGATGATGATTTGTTAGAGTTAGAGAAAGTGGAGAAGTT 885
Qy      278  ProValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGly 297
Db      886  CAAATTTCTATCACAAATGACATATCAATATCAATATTAATCGACATGTTCTCTCGGGA 945
Qy      298  SerGluThrThrSerLysAlaThrGluTrpValMetalGluLeuMetLysAsnProThr 317
Db      946  TCTGAACATCATCGACGACTATAATTTGGGCATTAGCTGAAATGATGAAGAACCAAGT 1005
Qy      318  GluLeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAsp 337
Db      1006  GTTTTAGCAAGGACACAGCTGAAGTAAAGCAAGCTTTGAAGGAGAAAAAGGTTTCAA 1065
Qy      338  GluSerArgPheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeu 357
Db      1066  CAGATTGATCTTGATGAGCTAAATATCTCAAGTTAGTTAATCAAGAANAACCTTAGAATG 1125
Qy      358  HisProProVal---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGly 376
Db      1126  CACCTCCAAATCTCTCTATTAGTTCTTAGAATGTTATGGAGGATACAAAGATTGATGTT 1185
Qy      377  TyrGluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspPro 396
Db      1186  TACAATATACCTTTCAAAACAAGAGTCATAGTTAATGCTGGCAATCGGACGAGATCCA 1245
Qy      397  AsnThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAsp 416
Db      1246  GAAAGTTGGGATGACCCCGAAAGCTTTATGCCAGAGATTTTGAGAAATGTTCTATTGAC 1305
Qy      417  TyrLysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGly 436
Db      1306  TTTCTTGGAAATCATCATCAGTTTATACCAATTTGGTCAGGAAGAAGGATTTGTCGGGA 1365
Qy      437  IleThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsn 456
Db      1366  ATGCTATTGTTGGTGTAGCTAAATGTTGGCAACCTTTTGTTCAGTCTACTTTATCCTTCGAT 1425
Qy      457  TrpGluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGly 476

```

Qy 167 AlaGlySerProValAsnIleThrIysIleIleTyrGlyIleIleIleSerIleMetIle 186

Search completed: May 31, 2006, 09:03:53
Job time : 8518 secs

This Page Blank (uspto)

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: May 31, 2006, 06:39:28 ; Search time 6894 Seconds
(without alignments)
6083.478 Million cell updates/sec

Title: US-10-759-813-2
Perfect score: 2574
Sequence: 1 MEQKNLSPSILFLVLI.....KIDKLPIPYQVSLGSNIS 500

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0
Searched: 48236798 seqs, 27959665780 residues
Total number of hits satisfying chosen parameters: 96473596
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=abs/ABSSWEB spool/US10759813/runat_31052006_060306_2636/app query.fasta_1
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abes02h
-USER=US10759813 @CNC 1.1 6323 @runat_31052006_060306_2636 -NCPU=6 -ICPU=3
-NO WMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est3:*
3: gb_est4:*
4: gb_est5:*
5: gb_est6:*
6: gb_hic:*
7: gb_est2:*
8: gb_est7:*
9: gb_est8:*
10: gb_est9:*
11: gb_gse1:*
12: gb_gse2:*
13: gb_gse3:*
14: gb_gse4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	954.5	37.1	1548	13	CL965804
2	947.5	36.8	1569	13	CL965067
3	947	36.8	1833	6	AY104522
4	934	36.3	1560	13	CL965799

5	927.5	36.0	1536	13	CL965911	CL965911
6	927	36.0	1575	13	CL977750	CL977750
7	923.5	35.9	1542	13	CL966810	CL966810
8	913.5	35.5	1557	13	CL966924	CL966924
9	910	35.4	1536	13	CL962214	CL962214
10	909.5	35.3	1581	13	CL967011	CL967011
11	907	35.2	1530	13	CL965809	CL965809
12	905	35.2	2259	13	CL965071	CL965071
13	901	35.0	1518	13	CL965068	CL965068
14	898	34.9	1539	13	CL965064	CL965064
15	897	34.8	1557	13	CL958628	CL958628
16	895	34.8	1572	13	CL965801	CL965801
17	890.5	34.6	1527	13	CL964411	CL964411
18	890.5	34.6	1536	13	CL968663	CL968663
19	888	34.5	1542	13	CL981555	CL981555
20	885	34.4	1593	13	CL960539	CL960539
21	881.5	34.2	1545	13	CL965805	CL965805
22	871	33.8	1539	13	CL958604	CL958604
23	856	33.3	1572	13	CL977749	CL977749
24	855	33.2	1572	13	CL965811	CL965811
25	849	33.0	1518	13	CL977242	CL977242
26	843.5	32.8	1560	13	CL965062	CL965062
27	841.5	32.7	1566	13	CL947270	CL947270
28	841.5	32.7	1566	13	CL965806	CL965806
29	839.5	32.6	1629	6	CNSOABEN	CNSOABEN
30	839	32.6	1902	13	CL965800	CL965800
31	836	32.5	732	10	DV153215	DV153215
32	831.5	32.3	1518	13	CL964423	CL964423
33	830	32.2	1362	13	CL965069	CL965069
34	814	31.6	1614	13	CL958110	CL958110
35	812.5	31.6	1422	13	CL965810	CL965810
36	811	31.5	1569	13	CL979405	CL979405
37	807.5	31.4	1656	6	CNSOAG90T	CNSOAG90T
38	803.5	31.2	1565	6	CNSOAG85	CNSOAG85
39	803	31.2	1580	6	CNSOABYN	CNSOABYN
40	800.5	31.1	1643	6	CNSOAB1J	CNSOAB1J
41	800	31.1	1479	13	CL965070	CL965070
42	796.5	30.9	1532	6	CNSOAG1J	CNSOAG1J
43	795	30.9	1602	13	CL958109	CL958109
44	785	30.5	1401	13	CL965803	CL965803
45	782	30.4	1584	13	CL963214	CL963214

ALIGNMENTS

RESULT 1
CL965804
LOCUS
DEFINITION
CL965804
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CL965804
OsIFCC012776 Oryza sativa Express Library Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.
CL965804
GSS.
Oryza sativa (indica cultivar-group)
Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP clade; Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 1548)
Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M., Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L., Wong, G.K.S., Deng, X.W. and Wang, J.
An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis
Unpublished (2004)
Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.

Db	820	GACGGCCCATGACGACGGGAGGAGGACCTCGTTGACGTGCTCTTCAGGATACAGAAA	879
Qy	274	AsnGlyAsnValProValProValThrAsnGluSerIleLysAlaSerValLeuGlnMet	293
Db	880	GATGGTGGCTCGAGGTGCTCTCACCATTTGGCAACATCAAGCAATCATCTCTGGACCTC	939
Qy	294	PheThrAlaGlySerGluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMet	313
Db	940	TTCAATGCAGGAGTGCAGACATCAGCAAAATACGCTCCAGTGGGTATGTGCAGAGCTAATG	999
Qy	314	LysAsnProThrCluLeuArgLysAlaGlnGluValArgGlnValPheGlyGluMet	333
Db	1000	AGGAACCCAAAAGTATGAGAAAAGCACACGAGAGTTAAGGAACCAACCTACCAAGGGAAG	1059
Qy	334	GlyLysValAspGluSerArgPheHisAspLeuLysPhePheLysLeuValValLysGlu	353
Db	1060	ACAACGGTACCGAAGCAGTCTCACAAATCTAAAGTACCTGAAGCTTGTTCATCAGGAG	1119
Qy	354	ThrLeuArgLeuHisProVal---ValLeuLeuProArgGluCysArgGluThrThr	372
Db	1120	ACGCTTAGGTATCATCCCGTGCCTTACTTCTCCCAAGAGAGTGCCCAAGGAGCATGC	1179
Qy	373	ArgIleAspGlyTyrrGluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIle	392
Db	1180	AATGTTCATCGGTACACGCTCCCAATACACTACCGCTTTTATAACGTGTGGGCAATC	1239
Qy	393	GlyArgAspProAsnThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAsp	412
Db	1240	AACAGGACCCCAAAATATGGGACATGGCCGAGATGTTTAAAGCCAGAACGCTTGACAAT	1299
Qy	413	CysAlaIleAspTyrrLysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArg	432
Db	1300	AGCATGATTGACTTCAAAGGTACAGACTTCGAGTTCGTGCAVTTGGGGTGGAAGG	1359
Qy	433	IleCysProGlyIleThrSerAlaIleThrAsnLeuGluTyrrValIleIleAsnLeuLeu	452
Db	1360	ATCTGTCCTGGCATGACATTTGACAGACTCCAAATATGGAGCTTGTGCTGTACACTCCTA	1419
Qy	453	TyrHisPheAsnTrpGluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGlu	472
Db	1420	TACACATTGACCTGGAGGCTTCTAGTGTATGTCGCTGGAAGAGTTAGATATGACTGAG	1479
Qy	473	AlaIleGlyAlaLeuArgLysLysIleAspLeuLysLeuLeuPro 488	
Db	1480	GACATGGGACTATCTGCCGAAGGAAGATGATCTTTACCTACATCCA 1527	
RESULT 2			
CL965067			
LOCUS		1569 bp DNA linear	GSS 21-SEP-2004
DEFINITION		OsIFC011616 Oryza sativa Express Library Oryza sativa (indica	
		cultivar-group) genomic, genomic survey sequence.	
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			

[illegible]


```

Db      1393  GAACCTGGCCCTGCCACCCCTTCTTACCACCTTCGATTGGAGCCTCCCTGCTGGGGTGGAAG 1452
Qy      465   ProGlnThrLeuAspMetThrGluAlaIleGlyGlyAlaLeuArgLysIleAspLeu 484
Db      1453  CCAAGTAGCTAGATATGACCGAGGAGATGGGCATCACTGTTAGGAGGAGNAGATGACCTA 1512
Qy      485   LysLeu 486
Db      1513  TACCTA 1518

RESULT 5
CL965911
LOCUS   CL965911                1536 bp      DNA      linear      GSS 21-SEP-2004
DEFINITION  OsIFCC013000 Oryza sativa Express Library Oryza sativa (indica
ACCESSION  CL965911
VERSION    CL965911.1  GI:52386501
KEYWORDS   GSS.
SOURCE     Oryza sativa (indica cultivar-group)
ORGANISM   Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP
clade; Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE  1 (bases 1 to 1536)
AUTHORS    Ma,L., Wang,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M.,
Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L.,
Wong,G.K.S., Deng,X.W. and Wang,J.
TITLE      An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
JOURNAL    Unpublished (2004)
COMMENT    Contact: Chen Chen
            Department of Bioinformatic
            Beijing Institute of Genomics
            Chinese Academy of Sciences, Beijing 101300, China
            Tel: 86-10-80481559
            Fax: 86-10-80488876
            Email: chenchen@genomics.org.cn
            Rice genomic sequence.
            Class: exon-trapped.
FEATURES   Location/Qualifiers
            source          1..1536
                        /organism="Oryza sativa (indica cultivar-group)"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:39946"
                        /clone_lib="Oryza sativa Express Library"
                        /note="Oryza sativa exon trapped genomic sequences "
ORIGIN
Alignment Scores:
Pred. No.:      2,036-95      Length:      1536
Score:          927.50      Matches:    199
Percent Similarity: 58.3%      Conservative: 100
Best Local Similarity: 38.8%      Mismatches: 181
Query Match:     36.0%      Indels:     33
DB:              13          Gaps:         12

US-10-759-813-2 (1-500) x CL965911 (1-1536)
Qy      6   LeuSerPheProSerIleLeuIleSerPheLeuLeuValLeuIleLeu----- 21
Db      10  CTCCTTGTAACAGCGCTCTCTCTCCGGGCTCGCGGCTCGCGGTGCTCCAGTCTGTCAG 69
Qy      22  ValValMetArgLeuTrpLysLysGln-----AsnProProProGly 36
Db      70  CTCGCCGTCGTCAACAGAGGGAAGAAGCAAGCGCGCGGTGTCACCGACGCCCGCGGG 129
Qy      37  ProTrpLysPheProIleIleGlyAsnLeuProHisLeuLeuLeuThrSerAspLeuGly 56
Db      130  CCATGGAGGCTCGCGGTGATCGGAGCATGCAACAC-----CTCGCCGGGAGAGTGGCG 183
Qy      57  HisGluArgPheArgAlaLeuAlaGlnIleTrpGlyProValMetSerLeuGlnIleGly 76
Db      184  CACCGCGCGCTGAGGAGCCTCGCGCGCTGCGAGCGCGCGCTGATGCTCCAGCTCGGC 243

```

```

Qy      77   GlnValSerAlaValValIleSerSerAlaGluAlaAlaLysGluValMetLysThrGln 96
Db      244  GAGACGCGCTGGTGGTGGTTCGCTCCAGGAGAGTGGCGCGGAGGTGCTCGAGGACGAC 303
Qy      97   AlaAspAlaPheAlaGlnArgProIleValLeuAspAlaGlnIleValPheTrpAsnArg 116
Db      304  GACGCCAACTTCGCCACCCGCCCAAGGCTGTCGCGCGGAGGTGCTGCTCTACCGCGGC 363
Qy      117  LysAspValLeuPheAlaSerTyxGlyAspHisTrpArgGlnMetLysLysIleTrpIle 136
Db      364  GCCGACATCTCTCTCGCCGTCGCGGAGTACTGGCGCAAGCTCCGCGAGCTCTCGGCC 423
Qy      137  LeuGluPheLeuSerAlaLysValGlnSerSerArgLeuIleArgGluGluGluMet 156
Db      424  GCCGAGGTGCTGGGCGGCGAAGCGGCTCTCTCTCCGCGCATCAGGAAACAGAGATG 483
Qy      157  GluAspAlaIleThrPheLeuArgSerLysAlaGlySer--ProValAsnIleThrLys 175
Db      484  GCGAGTCGGGTGGAAAGGATCCGTGCGGTCCGACCATCGGTCCCGTGATGTTAGTGA 543
Qy      176  IleIleTrpGlyIleIleIleSerIleMetIleArgThrSerValGlyAsnCysLysGln 195
Db      544  CTTTCTATGACATGGCAATCAGTATCGTCTCATGTGCATCTCTTTGGTAAG---AAACA 600
Qy      196  Lys-----GluArgLeuLeuSerValAlaAspAlaValAsnGluAlaAlaThrSerPhe 213
Db      601  AGGAATGTGACGAGTACCTTTTACGAATAAAGACTGGGATCAGTCTGGTAGTGGTTTC 660
Qy      214  GlyThrAlaAspAlaPheProThrTrpLys---LeuLeuHisTyrlleIleGlyAlaGlu 232
Db      661  AAGATCCCTGACCTCTTCCGACATGGAGAACGGTGCTTGTGCGAGTACGCGGCATGG 720
Qy      233  SerLysProArgArgLeuHisGlnGluIleAspAspIleLeuGluGluIleLeuAsnGlu 252
Db      721  CGTCCCTTGAATAATGTCATAGGATAGTGACCTCCACCTTTGGAAGAGGTCTACAGGAG 780
Qy      253  HisLysAla-----AsnLysProPheGluAlaAspAsn 263
Db      781  AGGAGAGCGCTGCAAGAGAAATGCAAGGCGCAGATTGGACATGGAGGACAAACGAGAA 840
Qy      264  LeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValProValProValThrAsn 283
Db      841  TTGGTGCATGTGCTCATCAAAATTCATGACGACCAAGGTGGC-----CACCTAGTAGA 891
Qy      284  GluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySerGluThrThrSerLys 303
Db      892  AATAGCATCAAGTCGGTGATATTTGATATGTTTACAGCGCGGAACAGGCACGTAGCATCA 951
Qy      304  AlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGluLeuArgLysAlaGln 323
Db      952  TCGCTAAATTTGGGGTATGTCGAACTAATAGGAACCCCAAGGGTGTGACCAAGTGTCAA 1011
Qy      324  GluGluValArgGlnValPheGlyGluMetGlyLysValAspGluSer-----ArgPhe 341
Db      1012  GGTGAGATCCGAGAGCGGCTTTCACGGCAAGGCTACCGTAGGGAGGGGTGACATCCAGTG 1071
Qy      342  HisAspLeuLysPhePheLysLeuValLysGluThrLeuArgLeuHisProVal 361
Db      1072  AGCAACCTCTCTACCTACCTTAGGCTCTTTATCAAGAGAGACCTTCGGTTCACCCCTCAGTG 1131
Qy      362  ---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTrpGluIleHis 380
Db      1132  CCTCTCTTGTGCCCCGGGAGAGCATTCATGTCGAGGTCAATGGGTATATACGATCCCT 1191
Qy      381  ProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsnThrTrpSer 400
Db      1192  GCAAGGTCACTGATCTGTTGTAATGCAATGGGCTATTGGAGAGAGACCCAAAGTATTGGAT 1251
Qy      401  GluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTrpLysGlyThr 420
Db      1252  GATCCTGAGGAATTCAGGCCAGACGATTTGAGGGCAACAAGGTGGACTTTTCAGGGACT 1311

```



```

Db      1189  ATGGGCTACGACATACCGAAGACACAACACTGTATTGTGAATGCGATGGCGCAATATGCAGA 1248
Qy      AspProAsnThrTrpSerGluProGlyLysAsnProGluArgPheLysAspCysAla 414
Db      1249  GACCCAAATATTGGATGATCCCGAAGATTCAACCCAGAGAGTTTGAGAACAAAGT 1308
Qy      IleAspTyrLysGlyThrPheGluLeuValProPheGlyAlaGlyLysArgIleCys 434
Db      1309  ATCGACTATAAGGAAGTAACCTTGAATTCCTTCATTTGGATCCGCGCGGTAGATGTGT 1368
Qy      ProGlyIleThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHis 454
Db      1369  GCGGCTATGAACCTTGGCATTTGCGAATGTAGAGCTCCCTTAGCTAGTAGCTTCTGTATCAT 1428
Qy      PheAsnTrpGluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIle 474
Db      1429  TTTGATTGGAAGTTGCTGATGGAATGATGCTGAAGATGTTGACATCAGATGCACCT 1488
Qy      GlyGlyAlaLeuArgLysLysIleAspLysLeuIleProIle 489
Db      1489  GGAATACTTGTAGGTAAACGCAATAGCCTAATTATGTGCCCGGTT 1533

```

RESULT 7

```

CL966810      1542 bp      DNA      linear      GSS 21-SEP-2004
LOCUS      OSIFCC014149 Oryza sativa Express Library Oryza sativa (indica
DEFINITION      cultivar-group) genomic, genomic survey sequence.
ACCESSION      CL966810
VERSION      CL966810.1 GI:52388272
KEYWORDS      GSS.
SOURCE      Oryza sativa (indica cultivar-group)
ORGANISM      Oryza sativa (indica cultivar-group)

```

```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP
clade; Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 1542)
Ma,L., Wang,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M.,
Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L.,
Wong,G.K.S., Deng,X.W. and Wang,J.
An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
Unpublished (2004)
Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488876
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.
Location/Qualifiers

```

```

FEATURES
source
1..1542
/organism="Oryza sativa (indica cultivar-group)"
/mo_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Express Library"
/note="Oryza sativa exon trapped genomic sequences"

```

```

ORIGIN
Alignment Scores:
Pred. No.:      5,918-95      Length:      1542
Score:      923.50      Matches:      191
Percent Similarity:      57.3%      Conservative:      99
Best Local Similarity:      37.7%      Mismatches:      197
Query Match:      35.9%      Indels:      19
Db:      13      Gaps:      6

```

```

US-10-759-813-2 (1-500) x CL966810 (1-1542)

```

```

Qy      1 MetGluGlnLysAsnLeuSerPheProSerIleLeuIleSerPheLeuLeuValLeu 20
Db      1 ATGGAGGACACAGCATACTACTA---TCTCTGGCCTTGTCCATGCTCTCCGCCTC 57

```

```

Qy      21 LeuValValMetArgLeuTrpLysLysGln-----AsnProProProGlyProTrp 38
Db      58 CTCTCCAAGCTGGTCTCGATTCCAAAGAACCCAGGCTGAACCTTCCCGCGGCGCATGG 117
Qy      39 LysPheProIleIleGlyAsnLeuProHisLeuLeuThrSerAspLeuGlyHisGlu 58
Db      118 ACGTTCGCGGTGATCGGCAGCATCCATCTCGCCAGCAACCCGACACC---CACAGG 174
Qy      59 ArgPheArgAlaLeuAlaGlnIleTyrGlyProValMetSerLeuGlnIleGlyVal 78
Db      175 GCGTTCGCGGCTCTGTCGACAGACACGCGGCCCTGATGCAGCTGTGCTGGCGAGGTG 234
Qy      79 SerAlaValValIleSerSerAlaGluAlaLysGluValMetLysThrGlnAlaAsp 98
Db      235 CCCGCGGTGGTGGGTTCAGCCGCGGAGCGCGGGAGATCCTCAGAAACACGAGACCTC 294
Qy      99 AlaPheAlaGlnArgProIleValLeuAspAlaGlnIleValPheTyrAsnArgLysAsp 118
Db      295 AGGTTTCGCGGACCGCCAGCTCACCAGCACCTCGCGCAGCGTCTCTTCGACGCCACGCAC 354
Qy      119 ValLeuPheAlaSerTyrGlyAspHisTrpArgGlnMetLysLysIleTrpIleLeuGlu 138
Db      355 ATCTTCTTCTCCCGTACGCGAGCGGTGGCGCAGCTGCGCAAGCTGTGCACGCGAGAG 414
Qy      139 PheLeuSerAlaLysLysValGlnSerSerArgLeuIleArgGluGluMetGluAsp 158
Db      415 CTCCTGACGGCCACCCGCGGTGAGGTGCTTCAGCCCGCTCCGCGAGGAGGAGGTGGCGGC 474
Qy      159 AlaIleThrPheLeuArgSerLysAlaGlySerProValAsnIleThrLysIleIleTyr 178
Db      475 CTCGTGGCGGAGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 534
Qy      179 GlyIleIleIleSerIleMetIleArgThrSerVal---GlyAsnCysLysGlnLysGlu 197
Db      535 AGGCTCGTCAACGAGCTGCTGATGAGGTGCTCCGTGCGGCGCGCGGTGAGGTGACGGGAC 594
Qy      198 ArgLeuLeuSerValAlaAspAlaValAsnGluAlaIleThrSerPheGlyThrAlaAsp 217
Db      595 GAGTTCCTGGCGCGCTCCACGAGGCCAACACGAGTGCAGCTGTGCTCACCCTGCGCGGAC 654
Qy      218 AlaPheProThrTrpLysLeuHisTyrIleIleGlyAlaGluSerLysProArg 237
Db      655 CTCTTCCCGTCTCGCAGGCTCGCGCGGATGTGGCGCGCGCGCGCGCGCGCGCGCGCGCG 714
Qy      238 LeuHisGlnGluIleAspAspIleLeuGluGluLeuAsnGluHisLys-----254
Db      715 AGCCGGAAGAGGATAGAGCGGATCATAGACACATCGTCCGGGAGCAGCAGGAGGTACATG 774
Qy      255 -----AlaAsnLysProPheGluAlaAspAsnLeu 264
Db      775 GGGAGCGCGCGCGCGCGCGCGAGCAGCAGCTGTGCGAGCTGCGAGGAGGAAAGACTGCTTC 834
Qy      265 MetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValProValProValThrAsnGlu 284
Db      835 CTCACGCTCTCTGTAGGCTGCGAGAGGAGGCGCGCACCGCGATCCCGATCAAAATGAA 894
Qy      285 SerIleLysAlaSerValLeuGlnMetPheThrAlaGlySerGluThrThrSerLysAla 304
Db      895 ATCATCGTGTGCTCTCTGTTGACATGTTTTCAGGGGGGTAGCGAGACATCAGCCACAGTG 954
Qy      305 ThrGluTrpValMetAlaGluLeuMetLysAsnProThrGluLeuArgLysAlaGlnGlu 324
Db      955 ATGATCTCGATCATGGCAGAGCTTATTCGGTGGCCCAAGAGTAAATGACAAAGTACAGCC 1014
Qy      325 GluValArgGlnValPheGlyGluMetGlyLysValAspGluSerArgPheHisAspLeu 344
Db      1015 GAGGTGCGCCCAAGCCCTCCAAGGGAGGTGCACAGTCACGTGACGAGTATATTGTAGGCTA 1074
Qy      345 LysPhePheLysLeuValValLysGluThrLeuArgLeuHis---ProProValValLeu 363
Db      1075 AACTATCTAAATAATGTTGATCAAGGAGACCTTACGATTTGCAATTGCGCTCGCCCTCCTCA 1134

```



```

Db      1006  GCGAAGGCACAGCTGAGGTAGGACAGAGCTGTTGGGAAGACATAGAGATCACAGAGAAAT 1065
Qy      340  ArgPheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHisPro 359
Db      1066  GATATTGTTGGCTGAGCTACCTCAAGATGGTGATCAAGAGGCGCTTCGGACTCCATTCC 1125
Qy      360  ProValValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 378
Db      1126  CCGGCTCCCTTGGCTGAACCTCGTAAATGTCGTGAGACAAACCAAGTCATAGGCTATGAT 1185
Qy      379  IleHisProAsnThrArgIleValValAenAlaTTPAlaIleGlyArgAspProAsnThr 398
Db      1186  ATACCAAAAGGCACATCTGTGTTGTGAATATGTGGCGGATATGTAGGAGCCCTAATTAC 1245
Qy      399  TrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyrLys 418
Db      1246  TGGGAGGATCCGAGAGTTTAAAGCCAGAGGTTTGAATAATATGTAGATTTTAAAG 1305
Qy      419  GlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIleThr 438
Db      1306  GGTAAATAACTTTGAATTCCTTCCATTTGGCTCTGGACGTAGAAATTTGTCGGGTATAAAC 1365
Qy      439  SerAlaIleThrAsnLeuGluTyrValIleIleAenLeuLeuTyrHisPheAsnTrpGlu 458
Db      1366  CTTGGGCTAGCCAACTTGAGCTAGCATTAGCTAGTCTTCTACCATTTTGTACTGGAAG 1425
Qy      459  LeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyAlaLeu 478
Db      1426  CTACCAATGAATGTTGCCCAAGGACCTTGACATGCAGGAGACCCCGGAATAGTTGCA 1485
Qy      479  ArgLysLysIleAspLeuLysLeuIleProIle 489
Db      1486  GCTAAACTAACTACTCTAAATATATGTGCCCTGTT 1518

RESULT 9
CL962214
LOCUS      OsIPCC007387 Oryza sativa Express Library Oryza sativa (indica
DEFINITION cultivar-group) genomic, genomic survey sequence.
ACCESSION CL962214
VERSION    CL962214.1 GI:52379162
KEYWORDS   GSS.
SOURCE      Oryza sativa (indica cultivar-group)
ORGANISM    Oryza sativa (indica cultivar-group)
REFERENCE   1 (bases 1 to 1536)
AUTHORS     Ma,L., Wang,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M.,
            Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L.,
            Wong,G.K.S., Deng,X.W. and Wang,J.
TITLE       An analysis of transcriptional regulation of the rice genome and
            its comparison to Arabidopsis
JOURNAL     Unpublished (2004)
COMMENT     Contact: Chen Chen
            Department of Bioinformatic
            Beijing Institute of Genomics
            Chinese Academy of Sciences, Beijing 101300, China
            Tel: 86-10-80481559
            Fax: 86-10-80488676
            Email: chenchen@genomics.org.cn
            Rice genomic sequence.
            Class: exon-trapped.
FEATURES
            Location/Qualifiers
            source          1..1536
                        /organism="Oryza sativa (indica cultivar-group)"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:39946"
                        /clone_lib="Oryza sativa Express Library"
                        /note="Oryza sativa exon trapped genomic sequences "
ORIGIN
Alignment Scores:

```

```

Pred. No.:      2,13e-93      Length:      1536
Score:          910.00      Matches:      195
Percent Similarity: 57.0%      Conservative: 91
Best Local Similarity: 38.8%      Mismatches: 184
Query Match:    35.4%      Indels:      32
Db:             13      Gaps:      8

US-10-759-813-2 (1-500) x CL962214 (1-1536)

Qy      12  LeuIleSerPheLeuLeuValLeuIleLeuValVal----- 23
Db      31  TTGATATCTTCTTCTCTCTCTCAAGCTCATCGCCGCTATGCTCGCCGTCGCGCAGG 90
Qy      24  -----ValMetArgLeuTrpLysGlnAsnProProGlyProTrpLysPhePro 41
Db      91  GAGAGTGGCTGCGGCTG-----CGCCCGGCGCGCTCGAGCTGCGC 132
Qy      42  IleIleGlyAenLeuProHisLeuLeuLeuThr-----SerAspLeuGlyHisGluArg 59
Db      133  CTATATCGGACGCTGACCACTCTCTCTGTACGGTACGGCGACCTGCCGACCGCGGCC 192
Qy      60  PheArgAlaLeuAlaGlnIleTyrGlyProValMetSerLeuGlnIleGlyGlnValSer 79
Db      193  ATGCGCGAGCTGCTCCCTGACCTACGGCCCTCATGCTCTCTCGGCTCGGCGCGTGGCC 252
Qy      80  AlaValValIleSerSerAlaGluAlaLysGluValMetLysThrGlnAlaAspAla 99
Db      253  ACGCTGGTGTCTCTCTCGCGAGCGCGCGGAGGTTCATGAGGCGCGCACGACCGCCGC 312
Qy      100  PheAlaGlnArgProIleValLeuAspAlaGlnIleValPheTyrAsnArgLysAspVal 119
Db      313  TTCGCGGCGCCACCTCAGCGCCACCATCGACATCTCTACGTGCGGCGGCGGACGACATC 372
Qy      120  LeuPheAlaSerTyrGlyAspHisTrpArgGlnMetLysLysIleTrpIleLeuGluPhe 139
Db      373  ATCTTCGGCGCCTACACCGAGCGGTGGCGGAGCTCCGCAAGGTGTGCGGCTCGAGCTC 432
Qy      140  LeuSerAlaLysLysValGlnSerSerArgLeuIleArgGluGluGluMetGluAspAla 159
Db      433  TTCAACACCGCGCGTCTCTCTCTCGCCCGCTCCTCGCGCGCGCGAGGAGGTTGGG----- 486
Qy      160  IleThrPheLeuArgSerLysAlaGlySerProVal-----Aen 172
Db      487  ---CGACTCTCTCCGCTCTCTCGCCGCTTCGCGGAGGCGCGCGCGCTGCTTCAAC 543
Qy      173  IleThrLysIleIleTyrGlyIleIleSerIleMetIleArgThrSerValGly--- 191
Db      544  CTCAGCGAGAGGATCTGCGCGCATGACCAACAGACTCCGTGTGTGCGCGCGGTCGCGCG 603
Qy      192  AsnCysLysGlnLysGluArgLeuLeuSerValAlaAspAlaValAsnGluAlaThr 211
Db      604  CGGTGCGACCAACCGCGGAGGTTCTCTGACGAGCTCGACAAGCGCGGTGCGGCTCACCGGC 663
Qy      212  SerPheGlyThrAlaAspAlaPheProTrpLysLeuLeuHisTyrIleIleGlyAla 231
Db      664  GGGATCAACTTGGCGGACCTGTACCCGCTCTCGCGGCTGTGTGCGCGGCTCAGCGCGGCC 723
Qy      232  GluSerLysProArgArgLeuHisGlnGluIleAspAspIleLeuGluGluIleLeuAen 251
Db      724  ACGCGGACATGGCGCGTTCGACAGAGAAACATCTACCGCATCGCGAGAGCATCATCCGC 783
Qy      252  GluHisLysAlaAenLysProPheGluAlaAsp-----AsnLeuMetAspValLeuLeu 269
Db      784  GACCGCGACGCGCGCGCGCGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 843
Qy      270  AsnLeuGlnLysAenGlyAsnValProValProValThrAsnGluSerIleLysAlaSer 289
Db      844  CGGCTGCGAGAGGCGCGCGCGCTCAAGTTCGCACTCACCCAGAGATCATCAGCAGCGCTC 903
Qy      290  ValLeuGlnMetPheThrAlaGlySerGluThrThrSerLysAlaThrGluTrpValMet 309
Db      904  ATTTTCGCATCTTCTCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 963

```

```
Qy 310 AlaGluLeuMetLysAsnProThrGluLeuArgLysAlaGlnGluValArgGlnVal 329
Db 964 TCGAGCTGATGAAGACCGCGATCTTGGCGAAGCGCGCAATCAGAGGTGAGGGAGACG 1023
Qy 330 PheGlyGluMetGlyLysValAspGluSerArgPheHisAspLeuLysPheLysLeu 349
Db 1024 TTCAAGGGACAAGACCTAACCGAGGATGATGTGCGCCAGCTGACTACCTGCAGCTG 1083
Qy 350 ValValLysGluThrLeuArgLysProProVal---ValLeuLeuProArgGluCys 368
Db 1084 GTGATCAAGAGACCTACGTTACACCCCGCAGCCCACTCTGTATCCCGAGGAGTGC 1143
Qy 369 ArgGluThrThrArgLysAspGlyTyrGluLeuHisProAsnThrArgLysValValAsn 388
Db 1144 CGCGAGACGTGCGAGGTGATGATGTGCGCAAGGGCACCAAGGTCTTTGTCAAC 1203
Qy 389 AlaTTPAlaIleGlyArgAspProAsnThrTTPSerGluProGlyLysPheAsnProGlu 408
Db 1204 GTGTGAAGATAGCGAGGGAAGCGAATATTGGGCGCATGGCGAGATATTTAGGCCAGAG 1263
Qy 409 ArgPheLysAspCysAlaIleAspTyrLysGlyThrThrPheGluLeuValProPheGly 428
Db 1264 AGATTTGAGATAGTACACTGATTTTAGGGGCGAGACTTTGAGTTCACTCCCGTTTGA 1323
Qy 429 AlaGlyLysArgLysCysProGlyIleThrSerAlaIleThrAsnLeuGluTyrValIle 448
Db 1324 GCGGGCGGAGGATGTCTGTATAGCGTAGGTAGTTAGCTAATATGGAGCTTGCTCTT 1383
Qy 449 IleAsnLeuLeuTyrHisPheAsnThrGluLeuAlaAspGlyIleThrProGlnThrLeu 468
Db 1384 GTAGCTCTCTTTACCACCTTGTATGGGAGCTTCCCGCGGCATCAATCTGAGGAGCTA 1443
Qy 469 AspMetThrGluAlaIleGlyGlyAlaLeuArgLysLysIleAspLeuLysLeuIlePro 488
Db 1444 GACATGACAGAGTATTTGGTATACAGGTAAAGGAAGTCCAAGCTTTTGTTCACGCA 1503
Qy 489 IlePro 490
Db 1504 ATACCT 1509.

RESULT 10
CL967011
LOCUS
DEFINITION
Oryza sativa (indica cultivar-group)
1581 bp DNA linear GSS 21-SRP-2004
OsIFCC014585 Oryza sativa Express Library Oryza sativa (indica
cultivar-group) genomic, genomic survey sequence.
ACCESSION
CL967011
VERSION
CL967011.1 GI:52388670
SOURCE
Oryza sativa (indica cultivar-group)
Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP
Clade; Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 1581)
Ma, L., Wang, J., Chen, X., Su, N., Li, L., Wang, X., Cao, M.,
Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
Wong, G. K. S., Deng, X. W. and Wang, J.
An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
Unpublished (2004)
Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.
Location/Qualifiers
1..1581
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
```

```
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Express Library"
/note="Oryza sativa exon trapped genomic sequences"
```

ORIGIN

```
Alignment Scores:
Pred. No.: 2,55e-93 Length: 1581
Score: 909.50 Matches: 211
Percent Similarity: 57.4% Conservative: 93
Best Local Similarity: 39.8% Mismatches: 181
Query Match: 35.3% Indels: 45
DB: 13 Gaps: 12

US-10-759-813-2 (1-500) x CL967011 (1-1581)

Qy 1 MetGluGlnLysAsnLeuSerPheProSerIleLeuIleSerPheLeuLeuValLeuIle 20
Db 1 ATGGAGCAA-----GTGCGTACTGTTTCATCTGCGCATGCGCCTTCTTGGCTCTCGTG 54
Qy 21 LeuValVal-----ValMetArgLeuTrpLysLysGlnAsn 32
Db 55 CTCCTCGTCAGGCTGGGCGCGCGCGCGCGCGCGCTGTCGCGCTC----- 99
Qy 33 ProProGlyProTrpLysPheProIleIleGlyAsnLeuProHisLeuLeuLeuThr 52
Db 100 ---CGCGCGGCGCGTGGCGGCTGCGGTCGTGCGCAACTCCACCAGCTGATGCTCCGG 156
Qy 53 SerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIleTyr-----GlyProVal 70
Db 157 GGACCGCTCGTCCACCGCACCATGGCGGACCTAGCGCGCGGCTCGACGACGCGCGCTC 216
Qy 71 MetSerLeuGlnIleGlyGlnValSerAlaValIleSerSerAlaGluAlaAlaLys 90
Db 217 ATGCGCTCAGCTCGGCGGCTGCCGTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 276
Qy 91 GluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleValLeuAspAlaGln 110
Db 277 GAGGTGACGCGCACGACACTCGACTTCGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCAG 336
Qy 111 IleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAspHisTrpArgGln 130
Db 337 CGCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCAG 396
Qy 131 MetLysLysIleTrpIleLeuGluPheLeuSerAlaLysValGlnSerSerArgLeu 150
Db 397 CTCGCGAGGTGCTGCTGAGATGCTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 456
Qy 151 IleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLysAlaGlySer--- 169
Db 457 GTGCGTAGGAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 516
Qy 170 -----ProValAsnIleThrLysIleIleTyrGly 179
Db 517 CCGACCGCGCACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 576
Qy 180 IleIleIleSerIleMetIleArgThrSerValGlyAsnCys-----LysGlnLysGlu 197
Db 577 GCGTGGCGGACGCGACGATGCGCGCGTATCGGGGACAGGTTTCAGAGAGCGGGAGGAG 636
Qy 198 ArgLeuLeuSerValAlaAspAlaValAsnGluAlaAlaThrSerPheGlyThrAlaAsp 217
Db 637 TTCTCTGGAGTCGATCAGGAGGCGGTG---AGGAGCTTCACGGGGTTTACCGCTGGACGAC 693
Qy 218 AlaPheProThrTrpLysLeuLeuHisTyrIleIleGly-----AlaGluSer 233
Db 694 CTGTTCCCGTCTCGAGGCTCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 753
Qy 234 LysProArgArgLeuHisGlnGluIleAspAlleLeu-----GluGluIleLeuAsn 251
Db 754 ACATCCCGAAGGGCCACCGACTATGACAGCGCTTCGCGCGCATGACGCGCGCGCGCGCG 813
Qy 252 GluHisLysAlaAsnLysProPheGluAlaAsp-----AsnLeuMetAsp 266
Db 266
```



```
QY 238 LeuHieGlnLuleAspAspIleLeuGluGluLileLeuAsnGluHisLysAlaAsnLys 257
Db 733 CGCGCGCGCGGATGGCTTCATCGACACCATCATCCAGGACACCCAGGAGACGAGA 792
QY 258 ---ProPheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsn 276
Db 793 GCGCGCGCGGACAGGAGGAGCTTCGACGCTCTCTGAGGCTCCAGAAAGACATGGAC 852
QY 277 ValProValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAla 296
Db 853 TCCAGTATCCCTCACCACCATGAACATCAATCCATCTTGATCGACATGTTCCGGCG 912
QY 297 GlySerGluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnPro 316
Db 913 GCGAGCGAGACGTTCAGCGACGAGCTGCTGAGTGGCGATGCGGAGCTGATGCGTAACCG 972
QY 317 ThrGluLeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGly----- 334
Db 973 GCGGTGATGCGCGCGGAGGAGGAGTGGCGGAGAGCTCGCGTCCGCGGACGAC 1032
QY 335 LysValAspGluSerArgPheHisAspLeuLysPhePheLysLeuValLysGluThr 354
Db 1033 AGGTAACCGAGACACCTCCGAGCTCCACTACTCGGCTGTCTCAGGAGAGCG 1092
QY 355 LeuArgLeuHisProProVal---ValLeuIleProArgGluCysArgGluThrThrArg 373
Db 1093 CTCGGCTCCACCGCGGAGCGCTCTCTCGCGGAGTGGCGGCGCGTGCAGG 1152
QY 374 IleAspGlyTyrGluLileHisProAsnThrArgIleValValAsnAlaThrAlaIleGly 393
Db 1153 GTGTTCGGTACAGCTCGCGCGGACGATGCTGCTGTAACGCTGGCGATCGCG 1212
QY 394 ArgAspProAsnThrTrpSerGluProGly---LysPheAsnProGluArgPheLysAsp 412
Db 1213 CGGAGCGCGCGCTGGCGCGCGGAGGAGTTCCTCGCGGAGAGTTTCGAGCGG 1272
QY 413 CysAlaIleAspTyrLysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArg 432
Db 1273 TCGGAGAGGAGCTTCAGGGGGGCGGACTTCGAGCTCATCCGTTCCGCGCGGAGGAG 1332
QY 433 IleCysProGlyIleThrSerAlaIleThrAsnLeuLutryrValIleIleAsnLeuLeu 452
Db 1333 ATCTGCCCGGATGCGCTTCGGGCTGGCGGACGTCGAGCTCGCGTCCGCGCGCTGCTG 1392
QY 453 TyrHisPheAsnTrpGluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGlu 472
Db 1393 TTCCTTCGATGAGGCTGCGGAGGAGTGGCGCGGAGGAGTGGCGCGGAGTGCAGTGC 1452
QY 473 AlaIleGlyAlaLeuArgLysLysIleAspLeuLysLeu-----Ile 487
Db 1453 GCGCGCGGATCAGCGTGGCGCGCGCTTCGGACCTACTGTTGCTGCGCGTGGCGGTGC 1512
QY 488 ProIleProTyrGln 492
Db 1513 CCGGTGCCAGCGAG 1527

RESULT 12
CL965071
LOCUS
DEFINITION
OsIRF001624 Oryza sativa Expressed Sequence Tag (EST) Oryza sativa (indica
cultivar-group) genomic, genomic survey sequence.
ACCESSION
CL965071
VERSION
CL965071.1 GI:52384833
KEYWORDS
GSS.
SOURCE
Oryza sativa (indica cultivar-group)
Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP
clade; Ehrhartoideae; Oryzoae; Oryza.
1 (bases 1 to 2529)
REFERENCE
Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,
Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
Wong, G.K.S., Deng, X.W. and Wang, J.
```

```
TITLE
JOURNAL
COMMENT
An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
Unpublished (2004)
Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.
FEATURES
Location/Qualifiers
source
1..2259
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Expressed Library"
/note="Oryza sativa exon trapped genomic sequences "
```

ORIGIN

Alignment Scores:

Pred. No.:	1.45e-92	Length:	2259
Score:	905.00	Matches:	191
Percent Similarity:	57.9%	Conservative:	102
Best Local Similarity:	37.7%	Mismatches:	195
Query Match:	35.2%	Indels:	18
DB:	13	Gaps:	8

US-10-759-813-2 (1-500) x CL965071 (1-2259)

QY 1 MetGluGlnLysAsnLeuSerPheProSerIleLeuLeuSerPheLeuLeuValLeuIle 20

Db 730 CTGAACAGAGAAATGAGGAGCAGTATCTCTTGTGGCCACCTCTACTGATCTCTCTC 789

QY 21 LeuValValMetArgLeuTrpLysLys---GlnAsnProProProGlyProTrpLys 39

Db 790 GCCTTCGTCAAGCTCAGGCCACGCAACATGCGAGAGACCCGCTCCGGGCGCATCGCAG 849

QY 40 PheProIleIleGlyAsnLeuProHisLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 59

Db 850 CTGCGCGTATCGGAGCGCTGACACAC-----CTGCGCGCGCGCTTCTCTACCGCGCG 903

QY 60 PheArgAlaLeuAlaGlnIleTyrGlyProValMetSerLeuGlnIleGlyGlnValSer 79

Db 904 CTGCGTGACCTTGCCTCGCGCGCGCGCGAGCTCATGCTCTCGGCTCGCGGAGCTCC 963

QY 80 AlaValValIleSerSerAlaGluAlaLysGluValMetLysThrGlnAlaAspAla 99

Db 964 GTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1023

QY 100 PheAlaGlnArgProIleValLeuAspAlaGlnIleValPheTyrAsnArgLysAspVal 119

Db 1024 TTCGCCACGCGCGCGCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1083

QY 120 LeuPheAlaSerTyrGlyAspHisTrpArgGlnMetLysLysIleTrpIleLeuGluPhe 139

Db 1084 GCGTTCGCGCGCGCGAGGAGCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1143

QY 140 LeuSerAlaLysLysValGlnSerSerArgLeuIleArgGluGluGluMetGluAspAla 159

Db 1144 CTGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1203

QY 160 IleThrPheLeuArgSerLysAlaGlySerProValAsnIleThrLysIleIleTyrGly 179

Db 1204 CTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1263

QY 180 IleIleIleSerIleMetIleArgThrSerValGlyAsn---CysLysGlnLysGluArg 198

Db 1264 TACGTGACGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1323

QY 199 LeuLeuSerValAlaAspAlaValAsnGluAlaAlaThrSerPheGlyThrAlaAspAla 218


```
Db 496 GTCCGCCGCGCGCGCGACCGGCGGACCGTGGAGCTCGCGGCGCGCGCTGTGCGCGCTC 555
Qy 181 llelleSerlleMetlleArgThrSerValGly---AsnCyLeYsGlnLysGluArgLeu 199
Db 556 GTGGCGGAGTCCACCGTGGCGCGCGTGTGGCGAGCGGTGGCGGCGCTCGACGTGTTTC 615
Qy 200 LeuSerValAlaAspAlaValAenGluAlaThrSerPheGlyThrAlaAspAlaPhe 219
Db 616 CTCGCCAGCTCGACCGCGCCATCGAGTCTGCTGGCTCAACGTGGCGGACCTGTGG 675
Qy 220 ProThrTrpLysLeuHisTyrIleleGlyAlaGluSerLysProArgArgLeuHis 239
Db 676 CGGTCTGCTGCTCGCGCGCGTCTCGACGCGCGCGTGGCGCCAGCGGAGAGTGTGCT 735
Qy 240 GlnGluIleAspAspIleLeuGluLleLeuAenGluHisLysAlaAenLysProPhe 259
Db 736 GACACGATGTTGGGTCTCGACGCGCATCATCCAGCGGCACCTGGAGNAGACTGGCGGC 795
Qy 260 GluAlaAspAsnLeuMetAspValLeuLeuAenLysLysAenGlnValProVal 279
Db 796 GCGCGGAGACATTTAGCTGCTACTAGGATACATAGGAGGCGGCTGGAGTTC 855
Qy 280 ProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySerGlu 299
Db 856 CTTCTCGACATGAGCGCGTCAAAATGCGTCTGTTGATGATCATCAGCGGAGGCTCGGAG 915
Qy 300 ThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAenProThrGluLeu 319
Db 916 AGCTCGCGACACCGTGGGTGGCGTTCGCGGAGCTAATCCGGAACCCGCGCGCGATG 975
Qy 320 ArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGluSer 339
Db 976 AAGAAGGCGACGCGGAGTTCCGCTGACTTCGAGGCGCGCGCGCTGTGCAGAGAC 1035
Qy 340 ArgPhe-----HisAspLeuLysPheLysLeuValLysGluThrLeuArgLeu 357
Db 1036 GCGTCTGCGCGTGGCGAGCTCCGCTACCTTCGCCCTCGTCTGAGGAGACGCTCGCGCTG 1095
Qy 358 HisProProVal---ValleulleProArgGluCysArgGluThrThrArgileAspGly 376
Db 1096 CACCGCGCGTCCGCTGCTCCGCGCGAGTCCGCGGAGCGCGTGGCGGTGTCGCGC 1155
Qy 377 TyrGluIleHisProAsnThrArgileValValAenAlaTrpAlaIleLysArgAspPro 396
Db 1156 TAGCAGCTCCGCGCGCGCGAGTGTGTGTGAACGCTGCGGCGATCGCGCGACGAG 1215
Qy 397 AsnThrTrp-----SerGluProGlyLysPheAenProGluArgPheLysAsp----- 412
Db 1216 CGGTACTGCGCGCGCGTTCGCGGAGGAGTTCCGCGCGGAGCGGTTCGCGACGCGGAG 1275
Qy 413 -----CysAlaIleAspTyrLysGlyThrThrPheGluLeuValProPheGlyAlaGly 430
Db 1276 GCGCGCGCGCGGTGGACTTCAAGGCGCGCATTCGAGCTCTGCGGTTCGCGCGCGCGC 1335
Qy 431 LysArgIleCysProGlyIleThrSerAlaIleThrAsnLeuGluTyrVallelleAen 450
Db 1336 CGCGCGATGTGCGCGGATGCGGTTCGGGCTCGCAACACGTGGAGCTCCACTCGCCAGC 1395
Qy 451 LeuLeuTyrHisPheAsnTrpGluLeuAlaAspGlyIleThrProGlnThrLeuAspMet 470
Db 1396 GTGCTCTTCCATTGACATGGAGGCGCTCCGCTGGTGGCTGACCCGACCGAGTTCACATG 1455
Qy 471 ThrGluAlaIleGlyAlaLeuArgLysLysIleAspLeuLysLeulleProIle 489
Db 1456 ACAGAGGCGTTCGCGCATCACGCGCGCGCAAGGCCAACCTCTGCTCGTCCCAATC 1512

RESULT 15
CL958628 1557 bp DNA linear GSS 21-SEP-2004
LOCUS
DEFINITION
OsiFCC001807 Oryza sativa Express Library Oryza sativa (indica
cultivar-group) genomic, genomic survey sequence.
ACCESSION
CL958628
VERSION
CL958628.1 GI:52372005
```

```
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
ORIGIN
Alignment Scores:
Pred. No.: 6.93e-92 Length: 1557
Score: 897.00 Matches: 189
Percent Similarity: 58.4% Conservative: 104
Best Local Similarity: 37.6% Mismatches: 183
Query Match: 34.8% Indels: 26
DB: 13 Gaps: 8
US-10-759-813-2 (1-500) x CL958628 (1-1557)
Qy 11 llellelleSerPheLeuLeuValleulleLeuValValMetArgLeuTrpLysLys 30
Db 22 CTACTTCTCTTGTCTGCTGTCGCCCTCCCTTCCTCTGCTGACCGTGGTGGTCTC 81
Qy 31 GlnAsnPro-----ProProGlyProTrpTrpLysPheProIlelle 43
Db 82 CGGCATGCGCGCGCGCGCGCGTGTGCTCGTCCGCGCGCTGCTGTCATC 141
Qy 44 GlyAsnLeuProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeu 63
Db 142 GGCACCTTCCACAC-----GTGCGCGCGCGCTCCCGCACCGCGCATGCGGACCTC 195
Qy 64 AlaGlnIleTyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValle 83
Db 196 GCGCGCGCGCGCGCGCGCTATGCTGCTCGCCTCTCGAGCTCGCGCTGCTGTCGCC 255
Qy 84 SerSerAlaGluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArg 103
Db 256 TGCACCGCGGAGGCGCGCGAGGTGACCAAGACCCACGACTCGCTTCGCGACGCGG 315
Qy 104 ProfileValLeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSer 123
Db 316 CCCATCACCCCGACCGGGAAGTCTCATGGCGGACAGCGTCCGGCTGCTGTTCGGCCCC 375
Qy 124 TyrGlyAspHisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLys 143
Db 376 TACGCGGAGGTTGAGGAGCGCTCGCAGGATCTGCACCTCGAGCTCTTAAGCGCGCGA 435
Qy 144 LysValGlnSerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeu 163
Db 436 CGCGTCCGCTCTCTCCGCGCGTGGCGGAGGAGGTGCGCGCGGTGCTCGCGCGCGTG 495
```

```
GSS.
Oryza sativa (indica cultivar-group)
Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BBP
clade; Euhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 1557)
Ma,L., Wang,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M.,
Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L.,
Wong,G.K.S., Deng,X.W. and Wang,J.
An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
Unpublished (2004)
Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.
Location/Qualifiers
1..1557
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Express Library"
/note="Oryza sativa exon trapped genomic sequences "
```

Search completed: May 31, 2006, 08:36:43
Job time : 6934 secs

```
QY 164 ArgSerLysAlaGly-----SerProValAsnIle 173
Db   :|||
496 GCGCGCGCGCGCGGTGGCGGCTTGACAGACCCCGGGCGCGACGCGGTGAACCTG 555
QY 174 ThrLysIleIleTyrGlyIleIleIleSerIleMetIleArgThrSerValGlyAsn--- 192
Db   :|||
556 AGCGAGCGGATTTCGGCGGTACGTGGCGGACTCCGCGGTGCGCGCCGTCATCCGGAGCAGG 615
QY 193 CysLysGlnLysGluArgLeuLeuSerValAlaAspAlaValAsnGluAlaAlaThrSer 212
Db   :|||
616 TTCAAGAACCGCGCGCTTCCTGCGGATGCTGGAGCGGAGGATGAAGCTCTCCGCGCG 675
QY 213 PheGlyThrAlaAspAlaPheProThrTrpLysLeuLeuHisTyrIleIleGlyAlaGlu 232
Db   :|||
676 CAGTGTTCGCGGACCTCTCCGCTGCGCGCGCGCGATGCTCTCAGCCGGATGCC 735
QY 233 SerLysProArgArgLeuHisGlnGluIleAspIleLeuGluIleLeuAsnGlu 252
Db   :|||
736 CGCCGATGAAGCGGAGCGAGGAGATGATGATTTCATCGACGACATCTTCCAGGAG 795
QY 253 HisLysAlaAsnLysProPheGluAla-----AspAsnLeuMetAspValLeuLeuAsn 270
Db   :|||
796 CATCAGAGAGACGCGCGCGCGCGCGCGAGGAGACTGCTGATGCTCTCTCCGCG 855
QY 271 LeuGlnLysAsnGlyAsnValProValProValThrAsnGluSerIleLysAlaSerVal 290
Db   :|||
856 ATCCAGAGCCAAACAGAGCAATCCCGCCTCACCAACGACAAACATCAAGACCGGTGATC 915
QY 291 LeuGlnMetPheThrAlaGlySerGluThrThrSerLysAlaAlaThrGluTrpValMetAla 310
Db   :|||
916 ATCGACATGTTCGTGCGGAGACGCGAGACGCGCGCGCGCTGCTGCGAGTGGACCATGTGCG 975
QY 311 GluLeuMetLysAsnProThrGluLeuArgLysAlaGlnGluValArgGlnValPhe 330
Db   :|||
976 GAGCTGATGAGGACCTAGGTGATGCGCAGGCGCGAGGACGAGGTCCGACGCGCCCTC 1035
QY 331 GlyLeuMetGlyLys-----ValAspGluSerArgPheHisAspLeuLysPhePheLys 348
Db   :|||
1036 GCCATTGCGCGGAGCAGCGCGTCCAGGAGAGCTGCGCGATCTCCCTTACCTGCAC 1095
QY 349 LeuValValLysGluThrLeuArgLeuHisProProVal---ValLeuIleProArgGlu 367
Db   :|||
1096 CTCCTCATCAAGGAGTCGTCGCGGTGCACCGCGGTGACGATGCTGCTGCGGAGGAG 1155
QY 368 CysArgGluThrThrArgIleAspGlyTyrGluIleHisProAsnThrArgIleValVal 387
Db   :|||
1156 TGCCGGAGCCGTGCGCGGTGATGGCTTCGAGTCCCGGAGGCGGTGATGTGCTCGTC 1215
QY 388 AsnAlaTrpAlaIleGlyArgAspProAsnThrTrpSerGluProGlyLysPheAsnPro 407
Db   :|||
1216 AACGCGTGGCGATCGCGAGGACCCGCGCACTGGGACTCGCGGAGGAGTTGCGCGCG 1275
QY 408 GluArgPheLys---AspCysAlaIleAspTyrLysGlyThrThrPheGluLeuValPro 426
Db   :|||
1276 GAGAGTTCGAGGCGCGCGCGCGCGCGGACTTCAAGGGAGCGGACTTCGAGTACATACCG 1335
QY 427 PheGlyAlaGlyLysArgIleCysProGlyIleThrSerAlaIleThrAsnLeuGluTyr 446
Db   :|||
1336 TTCGGCGCGCGCGCGGATGCCCCGGGATGGCGTTCGGGCTGGCCAAACATGGAGCTC 1395
QY 447 ValIleIleAsnLeuLeuTyrHisPheAsnTrpGluLeuAlaAspGlyIleThrProGln 466
Db   :|||
1396 GCGCTCGCGCGCTCTCTGTACACTTCGACTGGGAGCTGCCCGCGGGATGCTTCCCGCG 1455
QY 467 ThrLeuAspMetThrGluAlaIleGlyGlyAlaLeuArgLysLysIleAspLeuLysLeu 486
Db   :|||
1456 GAGCTGGACATACGAGAGCGGTGGGGCTCACACGCGCGGTGCTCCGACCTCTCTGTT 1515
QY 487 IlePro 488
Db   :|||
1516 GTCCCT 1521
```

This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 31, 2006, 06:03:06 ; Search time 913 Seconds
(without alignments)
5727.477 Million cell updates/sec

Title: US-10-759-813-2

Perfect score: 2574

Sequence: 1 MEQKNLSPSILISFLVLI.....KIDKLPIPYQVSLGNSIS 500

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delopt 6.0, Delext 7.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlh
-Q/abs/ABSWEB spool/US10759813/runat_31052006_060302_2579/app query.fasta_1
-DB=N_Geneseq -QFM=facatp -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human4.0.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abes05h
-USER=US10759813 @CN 1 1 761 @runat 31052006_060302_2579 -NCPU=6 -ICPU=3
-NO WMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV -TIMEOUT=120
-WARN -TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_8.*

1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005s.*
15: Geneseqn2006s.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2574	100.0	1733	6	ABK11137
2	1164	45.2	1538	12	ADN10568 Nicotiana
3	1164	45.2	1538	14	ADZ65256 Tobacco f

4	1164	45.2	1538	14	ADZ65803	Adz65803 Tobacco c
5	1164	45.2	1538	14	AE06775	Ae06775 Tobacco p
6	1164	45.2	1538	15	AEF97168	Aef97168 Tobacco c
7	1149	44.6	1673	12	ADN10646	Adn10646 Nicotiana
8	1149	44.6	1673	12	ADN10626	Adn10626 Nicotiana
9	1149	44.6	1673	14	ADZ65334	Adz65334 Tobacco f
10	1149	44.6	1673	14	ADZ65314	Adz65314 Tobacco f
11	1149	44.6	1673	14	ADZ65861	Adz65861 Tobacco c
12	1149	44.6	1673	14	ADZ65881	Adz65881 Tobacco c
13	1149	44.6	1673	14	AE06853	Ae06853 Tobacco p
14	1149	44.6	1673	14	AE06833	Ae06833 Tobacco p
15	1149	44.6	1673	15	AEF97226	Aef97226 Tobacco c
16	1149	44.6	1673	15	AEF97226	Aef97226 Tobacco c
17	1148	44.6	1673	12	ADN10624	Adn10624 Nicotiana
18	1148	44.6	1673	14	ADZ65312	Adz65312 Tobacco f
19	1148	44.6	1673	14	ADZ65859	Adz65859 Tobacco c
20	1148	44.6	1673	14	AE06831	Ae06831 Tobacco p
21	1148	44.6	1673	15	AEF97224	Aef97224 Tobacco c
22	1144	44.4	1657	12	ADN10610	Adn10610 Nicotiana
23	1144	44.4	1657	14	ADZ65298	Adz65298 Tobacco f
24	1144	44.4	1657	14	ADZ65845	Adz65845 Tobacco c
25	1144	44.4	1657	14	AE06817	Ae06817 Tobacco p
26	1144	44.4	1657	15	AEF97210	Aef97210 Tobacco c
27	1142	44.4	1673	12	ADN10628	Adn10628 Nicotiana
28	1142	44.4	1673	14	ADZ65316	Adz65316 Tobacco f
29	1142	44.4	1673	14	ADZ65863	Adz65863 Tobacco c
30	1142	44.4	1673	14	AE06835	Ae06835 Tobacco p
31	1142	44.4	1673	15	AEF97228	Aef97228 Tobacco c
32	1131	43.9	1610	12	ADN10676	Adn10676 Nicotiana
33	1131	43.9	1610	12	ADN10612	Adn10612 Nicotiana
34	1131	43.9	1610	14	ADZ65300	Adz65300 Tobacco f
35	1131	43.9	1610	14	ADZ65364	Adz65364 Tobacco f
36	1131	43.9	1610	14	ADZ65911	Adz65911 Tobacco c
37	1131	43.9	1610	14	ADZ65847	Adz65847 Tobacco c
38	1131	43.9	1610	14	AE06819	Ae06819 Tobacco p
39	1131	43.9	1610	14	AE06883	Ae06883 Tobacco p
40	1131	43.9	1610	15	AEF97276	Aef97276 Tobacco c
41	1131	43.9	1610	15	AEF97212	Aef97212 Tobacco c
42	1129.5	43.9	1576	12	ADN10564	Adn10564 Nicotiana
43	1129.5	43.9	1576	14	ADZ65252	Adz65252 Tobacco f
44	1129.5	43.9	1576	14	ADZ65799	Adz65799 Tobacco c
45	1129.5	43.9	1576	14	AE06771	Ae06771 Tobacco p

ALIGNMENTS

RESULT 1
ABK11137
ID ABK11137 standard; cDNA; 1733 BP.
XX
AC ABK11137;
XX

DT 05-JUN-2002 (first entry)

XX cDNA encoding Euphorbia lagascae cytochrome P450 enzyme.

XX Cytochrome P450 enzyme; synthesis of delta 12-epoxy fatty acid; epoxide;
KW vernolic acid; modified fatty acid; oil; commercial crop; plasticiser;
KW crosslinking coating application; setting printing ink; transgenic;
KW plant; EST; clone eelc_pk002_14; gene; ss; expressed sequence tag.

XX Euphorbia lagascae.

XX Key Location/Qualifiers
FT CDS 14..1516
FT /*tag= a
FT /product= "Cytochrome P450 enzyme"

XX WO200208269-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US022790.

ID ADN10568 standard; cDNA; 1538 BP.
 XX ADN10568;
 XX AC
 XX DT
 XX 15-JUL-2004 (first entry)
 XX Nicotiana p450 enzyme encoding cDNA SEQ ID NO:153.
 DE plant; p450; enzyme; Nicotiana; plant growth regulant; gene therapy;
 XX phenotype; gene; ss.
 KW Nicotiana sp.
 XX OS
 XX WO2004035745-A2.
 XX PN
 XX PD
 XX 29-APR-2004.
 XX 16-OCT-2003; 2003WO-US032722.
 XX PF
 XX 16-OCT-2002; 2002US-0418933P.
 PR 08-JUL-2003; 2003US-0485368P.
 PR 18-SEP-2003; 2003US-0503989P.
 XX PA (USM-) US SMOKELESS TOBACCO CO.
 XX PI
 XX Xu D;
 XX WPI: 2004-348441/32.
 DR P-PSDB; ADN10569.
 XX PT
 XX New isolated nucleic acid molecules and encoded cytochrome P450 enzymes
 from Nicotiana plants, useful for altering plant phenotypes.
 XX Claim 1; SEQ ID NO 153; 198pp; English.
 XX PS
 XX The present sequence represents a plant p450 enzyme nucleic acid molecule
 isolated from Nicotiana. Also described: (i) an isolated p450 protein
 from Nicotiana; (2) a transgenic plant comprising the nucleic acid
 molecule described above; (3) a method of producing the transgenic plant,
 comprising: (i) operably linking the above nucleic acid molecule with a
 promoter functional in the plant to create a plant transformational
 vector; (ii) transforming the plant with the vector; (iii) selecting a
 plant cell transformed with the transformation vector; and (iv)
 regenerating a transformation plant from the transformed plant cell; and
 (4) a method of selecting a plant containing the above nucleic acid
 molecule, where the plant is analysed for the presence of the above
 nucleic acid sequences. The p450 sequences have plant growth regulant
 activity, and can be used in gene therapy. Compositions and methods from
 the present invention are useful for altering plant phenotypes.
 XX SQ
 XX Sequence 1538 BP; 464 A; 278 C; 334 G; 462 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 4,5e-102 Length: 1538
 Score: 1164.00 Matches: 226
 Percent Similarity: 63.4% Conservative: 88
 Best Local Similarity: 45.7% Mismatches: 165
 Query Match: 45.2% Indels: 16
 DB: 12 Gaps: 5

US-10-759-813-2 (1-500) x ADN10568 (1-1538)

QY 12 LeuileSerPheLeuLeuValLeuLeuValValMetArgLeuTrpLysGln 31
 Db 48 TTGGTTTCCATTTCCTATTTCTATCTTTCTCTTTTGTAAAGAAATGGAAGAACTCG 107
 QY 32 AsnPro-----ProProGlyProTrpLysPheProileleleGlyAsnLeu 46
 Db 108 AATAGCAAAAGGAAAAAATGCAACAGGTCCATGCAAACTACCAATACTAGGAAGTATG 167
 QY 47 ProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle 66
 Db 168 CTTTCATATGGTT-----GGTGGACTACCAACCCATGTCCTTTAGAGATTTAGCAAAAAA 221

QY 67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSerAla 86
 Db 222 TATGGACCCCTTATGCACCTTCATTTAGGTGAAGTTCTGACGTTTGGTATTCTCTCT 281
 QY 87 GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
 Db 282 GATATGGCAAAAGAGTACTAAAACTCATGACATCGCTTTTCGCGTCTAGGCTAGCCTT 341
 QY 107 LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
 Db 342 TTGGCCCCGGAGATTGCTGTTACAAATAGGTCTGATCTTTGCGGTTTCCCTATGGCGAT 401
 QY 127 HisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysValGln 146
 Db 402 TATTGGAGACAAATGCTAAATATATGTCTTGGAAAGTGTCTAGTGCACCAAGATGTCG 461
 QY 147 SerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLys 166
 Db 462 ACATATAGCTCTATTAGGCGCGATGAAGTTCTTCGTCTCTCTTAATTTATCCGTCATCT 521
 QY 167 AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleSerIleMetIle 186
 Db 522 TCTGTGAGCCCTGTTAATATTACGGAAGAGTCTTTTGTTCACAAAGCTCCATGACATGT 581
 QY 187 ArgThrSerValGlyAsnCys--LysGlnLysGluArgLeuLeuSerValAlaAspAla 205
 Db 582 AGATCAGCGTTTGGCGAGTATTCAAGGAGCAAGACAAATTTATACAACTAATTAAGAA 641
 QY 206 ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu 225
 Db 642 GTTATACTCTTAGCAGGAGGTTTGATGTGGCTGACATATTCCCTTCATACAAAGTCTCTT 701
 QY 226 HistyrIleIleGlyAlaGluSerLysProArgArgLeuHisGlnGluIleAspAspIle 245
 Db 702 CATGTCTCAGTGGAAATGAAGGTAAGATTATGAATGCACACCATCAAGGTAGTGTATT 761
 QY 246 LeuGluGluIleLeuAsnGluHis-----LysAlaAsnLysPro 258
 Db 762 GTTGAGATGTCATCAACGAGCACAGAAAAAATCTTGCAATGGGAAAAAATAATGAGCG 821
 QY 259 PheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValPro 278
 Db 822 TTAGGAGGTGAAGATTAAATCATGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 881
 QY 279 ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer 298
 Db 882 TTTCTCTATCACCACCAACCAACATCAAAAGCTATAATCTTTTGACATGTTTCTGCTGGA 941
 QY 299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
 Db 942 GAGACTTCATCTCAACAATTTGTGGGCTATGGTGGAAATGGTGGAAATAATCAACTGTA 1001
 QY 319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
 Db 1002 TTTGCGAAAGCTCAAGCAGAACTAAGAGATGATTCATTAGAGAAAAAGAAAACTTTGATGA 1061
 QY 339 SerArgPheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHis 358
 Db 1062 AATGATGTGGAGGAGCTAAACTATCTAAAGTTAGTCAATTAAGAACTCTTAAGACTTCAT 1121
 QY 359 ProProVal---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyr 377
 Db 1122 CCACCGGTTCCATTTTGTCTCCAGAGAGATGTAGGGAAGACAAATATAACGGCTAC 1181
 QY 378 GluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsn 397
 Db 1182 ACTATTTCTGTAAAGACCAAAAGTCATGGTTAATGTTGGGCATTTGGGAAGAGATCCAAA 1241
 QY 398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr 417
 Db 1242 TATTGGGATGATGCAAGAACTTTTAAAGCCAGAGATTTTGACGTCGTCTTAAGGATTTT 1301

QY 418 LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIle 437
 DB 1302 GTTGTAATAATTTTGAATATCTTCAATTTGGTGGGAAGGAGGATTTGCCAGGAT 1361
 QY 438 ThrSerAlaIleThrAsnLeuGluTrValIleAsnLeuLeuTyrHisPheAsnTrp 457
 DB 1362 TCGTTTGGTTAGCTAATGCTTATTTGCCATTGGCTCAATTACTTTATCACCCTTTGATTGG 1421
 QY 458 GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyAla 477
 DB 1422 GAACCTCCCACTGGGAATCAACCAAGCAGCTTGGACTTGACTGAGTTGGTGGAGTAACT 1481
 QY 478 LeuArgLysLysIleAspLysLeuIleProIleProTyrGln 492
 DB 1482 GCCGCTAGAAAAGTGACCTTACTTGGTTGGGACTCCTCTATCAA 1526

RESULT 3

ADZ65256

ID ADZ65256 standard; cDNA; 1538 BP.

XX

AC ADZ65256;

XX 14-JUL-2005 (first entry)

XX Tobacco full length cDNA for cytochrome p450 clone D90A-BB3.

DE Cytochrome p450; ss; secondary metabolite; ethylene; plant; senescence;
 KW normicotine; transgenic plant; gene.

XX Nicotiana tabacum.

XX WO2005038033-A2.

XX 28-APR-2005.

XX 15-OCT-2004; 2004WO-US034065.

XX 16-OCT-2003; 2003US-00686947.

XX 29-APR-2004; 2004US-05662359.

XX 03-SEP-2004; 2004US-00934944.

XX (USM-) US SMOKELESS TOBACCO CO.

XX Xu D;

XX WPI; 2005-315717/32.

XX P-PSDB; ADZ65257.

XX New nucleic acid molecule encoding cytochrome P450 enzymes in Nicotiana,
 PT useful in developing tobacco plants with altered phenotypes.
 XX Disclosure; SEQ ID NO 153; 226pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) from
 CC Nicotiana, where the nucleic acid molecule comprising any of the 59
 CC nucleic acid sequences of SEQ ID NOS: 299-357 (NOTE: The claims refer to
 CC SEQ ID NOS 299-357 as nucleic acids but these sequences (apart from SEQ
 CC ID NO 356) are all proteins and appear as ADZ65402-ADZ65460. The nucleic
 CC acids of the invention encode cytochrome p450 enzymes whose expression is
 CC induced by ethylene and/or plant senescence. Also included are a
 CC transgenic plant comprising (I), a method of producing a transgenic
 CC plant, a method of selecting a plant containing a nucleic acid molecule
 CC (where the plant is analyzed for the presence of nucleic acid sequence of
 CC ADZ65402-
 XX
 XX

SQ Sequence 1538 BP; 464 A; 278 C; 334 G; 462 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4,5e-102 Length: 1538
 Score: 1164.00 Matches: 226
 Percent Similarity: 63.4% Conservative: 88
 Best Local Similarity: 45.7% Mismatches: 165
 Query Match: 45.2% Indels: 16

DB: 14 Gaps: 5
 US-10-759-813-2 (1-500) x ADZ65256 (1-1538)
 QY 12 LeuIleSerPheLeuValIleLeuValValMetArgLeuTriPlysLysGln 31
 DB 48 TTGGTTTCCATTTTCCATTTCTATCTTTCTTTTGTAGGAATAGAAAGAACTCG 107
 QY 32 AsnPro-----ProProGlyProTriPlysPheProIleIleGlyAsnLeu 46
 DB 108 AATAGCCAAAGAAAAAATTCGCCACGAGTCCATGTAACCTACCAATACTAGGAAGTATG 167
 QY 47 ProHisLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle 66
 DB 168 CTTCATATGGTT-----GGTGGACTACCAACCATGTCCTTAGAGATTTTAGCCAAAAA 221
 QY 67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValIleSerSerAla 86
 DB 222 TATGACCGCTTATGACCTTCAATAGGTGAAGTTCTGCAGTTGTGGTACTTCTCTCT 281
 QY 87 GluAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
 DB 282 GATATGGCAAAAGAAAGTACTAAAAACTCATGACATCGCTTTCCGCTCTAGGCTAGCCTT 341
 QY 107 LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
 DB 342 TTGGCCCCGGAGATTGCTGTACAATAGGTCTGATCTTGGCTTTTGGCCCTTATGCGAT 401
 QY 127 HisTrpArgGlnMetLysLysIleTriPheLeuGluPheLeuSerAlaLysLysValGln 146
 DB 402 TATTTGGAGACAAATGCGTAAATATGTCTTTGGAGTGTCTCAGTCCCAAGAAATGTCGG 461
 QY 147 SerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLys 166
 DB 462 ACATATAGCTCTATTAGCGCGATGAAGTTCTTCGCTCTCTTAATTTTATCCGGTCATCT 521
 QY 167 AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleIleSerIleMetIle 186
 DB 522 TCTGTGTAGCCTGTTAATATTACGAAAGAGATCTTTTGTTCACAAGCTCCATGACATGT 581
 QY 187 ArgThrSerValGlyAsnCys---LysGlnLysGluArgLeuLeuSerValAlaAspAla 205
 DB 582 AGATCAGCGTTTGGCAAGTATTCAAGGAGCAAGCAAAATTTATACAACATAATTAAGAA 641
 QY 206 ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTriPlysLeuLeu 225
 DB 642 GTTATACTCTTAGCAGGAGGGTTGATGTGGCTGACATATTCCTTCCATACAAAGTCTCTT 701
 QY 226 HistyrlleIleGlyAlaGluSerLysProArgArgLeuHisGlnGluIleAspAspIle 245
 DB 702 CATGTGCTCAGTGGGAATGAAGGTAAATATGAATGACACACCATGAAGTAGATGCTATT 761
 QY 246 LeuGluGluIleLeuAsnGluHis-----LysAlaAsnLysPro 258
 DB 762 GTTGAGATGTCATCAAGCAGCACAAAGAAATCTTGCAATTTGGGAAACATAATGAGCG 821
 QY 259 PheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValPro 278
 DB 822 TTAGGAGGTGAAGATTAAATTGATGTTCTTCTTAAAACTTATGAATGATGAGGCGCTCAA 881
 QY 279 ValProValThrAsnGluSerIleLysValAserValLeuGlnMetPheThrAlaGlySer 298
 DB 882 TTTCTCTATCACCACGACCAACATCAAGCTATAATCTTTTGACATGTTTCTGCTGCTGGAACA 941
 QY 299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
 DB 942 GAGACTTTCATCGTCAACAATTTGTGGGCTATGGTGGAAATCGTGAAAAATCCAACTGTA 1001
 QY 319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
 DB 1002 TTTGGAAAGCTCAAGCAGAGAGTGAAGATGCAATTTAGAGAAAAAGAACTTTTGTGATGA 1061
 QY 339 SerArgPheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLysHis 358

Qy	226	HistyrillelleGlyAlaGluSerLysProArgArgLeuHisGlnGluIleAspAsePile	245
Dy	702	CATGTGTCAGTGAAGTAAGAAGGGTAAAGATTATTAAGATCACACCATTAAGGTAGATGCTATT	761
Qy	246	LeuGluGluIleLeuAsnGluHis-----LysAlaAsnLysPro	258
Dy	762	GTTGAGNAATGTCATCAACGAGCACAAGAAAATCTTGCAATTGGGAAAAACTNATGGAGCG	821
Qy	259	PheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValPro	278
Dy	822	TTAGGAGGTGAAGATTTAAATTGATGTTCTTCTATAAACCTTATGAATGATGAGGCCTTCAA	881
Qy	279	ValProValThrAsnGluSerIleLysLysAserValLeuGlnMetPheThrAlaGlySer	298
Dy	882	TTTCCTATCACCAACGACAACATCAAAAGCTATAATCTTTGCACATGTTTGCTGCTGGHACA	941
Qy	299	GluThrThrSertysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu	318
Dy	942	GAGACTTCATCGTCAACAATTTGTGTGGCTATGTTGGAATGTTGMAAAATCCAACCTGTA	1001
Qy	319	LeuArgLysAlaGlnGluValatqGlnValPheGlyGluMetGlyLysValaspGlu	338
Dy	1002	TTTCGAAAGCTCAAGCAGAAGTAAGAGATGCCATTTAGAGAAAAGAAAACCTTTTGATGAA	1061
Qy	339	SerArgPheHisAspLeuLysPhePheLysLeuValVallysGluThrLeuArgLeuHis	358
Dy	1062	AATGNTGTGGAGGAGCTAAACTATCTAAAGTTAGTCAATTAAGAAAACCTTAAGACTTCAT	1121
Qy	359	ProProVal---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyr	377
Dy	1122	CCACCGGTTCACCTTTTGCTCCCCAAGAGATGTAGGGAAGAGACAAATATAAACGGCGTAC	1181
Qy	378	GluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsn	397
Dy	1182	ACTATTCTGTAAAGACCACCAAGTCATGTTAAATGTTTGGCATTTGGGAAGAGATCCAAA	1241
Qy	398	ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr	417
Dy	1242	TATTGGGATATGCAGAAAACCTTTTAAACCCAGAGAGATTTGAGCAGTCTCTTAAGGATTTT	1301
Qy	418	LysglyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIle	437
Dy	1302	GTTGGTAAATTTTGAATATCTTCCATTTTGGTGGTGAAGGAGGATTTGTCACGGGATT	1361
Qy	438	ThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrp	457
Dy	1362	TCGTTTTGGTTTAGCTAATGCTTATTTGCCATTTGGCTCAANTTACTTTATCATCTTTGATTG	1421
Qy	458	GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAla	477
Dy	1422	GAATCCCCACTGGAAATCAACCAACGAGCTTGGACTTGGACTGGATTTGGTTGGNGTAAC	1481
Qy	478	LeuArgLysLysIleAspLeuLysLeuIleProIleProTyrGln	492
Dy	1482	GCCGCTAGAAAAAGTGACCTTTACTTTGTTGCCACTCTTATCAA	1526

RESULT 5

RESULT 5
AEE06775
ID AEE06775 standard; cDNA; 1538 BP.

AC AEE06775:

XX
DT 26-JAN-2006 (first entry)

XX DE FEB 1963

XXI

KW transgenic plant; senescence.

XX OS *Nicotiana tabacum.*

XX
PNI
W03006111217-12

[illegible]

24-NOV-2005. 27-APR-2005; 2005WO-US014803. 29-APR-2004; 2004US-0566235P. 03-SEP-2004; 2004US-0093494. 03-SEP-2004; 2004US-0607357P. 17-SEP-2004; 2004US-0094307. 15-OCT-2004; 2004WO-US034065. 15-OCT-2004; 2004WO-US034218. 25-JAN-2005; 2005US-0645764P. 24-MAR-2005; 2005US-0665097P. 24-MAR-2005; 2005US-0665451P. (US5M-) US SMOKELESS TOBACCO CO.

Xu D. Nielsen MT:

WPT: 2005-786788/80

P-PSDB; AEE06776.

Producing a tobacco plant having decreased expression of a nicotine demethylase gene comprises crossing a first tobacco plant with a second tobacco plant and germinating the collected seed of an F1 progeny plant.

Claim 32: SEO ID NO 162: 641pp: English.

The invention relates to a breeding method for producing a tobacco plant with reduced expression of a nicotine demethylase gene comprising crossing a first tobacco plant with variant nicotine demethylase gene expression with a second tobacco plant at least one phenotypic trait associated with reduced expression of a nicotine demethylase gene. The offspring produced by said cross are selected to produce a progeny plant the seed of which is used to produce a tobacco plant having reduced expression of a nicotine demethylase gene. Also included are methods for breeding a nicotine demethylase gene, also included are methods for breeding a tobacco seed, and a tobacco plant, both producing a tobacco seed.

deficiency trait into a tobacco plant, producing a tobacco seed, developing a tobacco plant in a tobacco breeding program, a tissue culture of regenerable tobacco cells obtained from the tobacco plant of the invention, producing a tobacco product, a breeding method for producing a tobacco plant having a modified attribute, a method of breeding an attribute into a tobacco plant, a tobacco plant or its components produced by the method of breeding a nicotine demethylase deficiency trait into a tobacco plant, producing tobacco seed, producing a tobacco plant having a modified attribute or developing a tobacco plant in a tobacco breeding program, an isolated genetic marker comprising a nucleic acid sequence that is substantially identical to a nucleic acid sequence given in the specification (the nucleic acids comprise isolated cytochrome p450 cDNAs), an expression vector comprising the nucleic acid sequence, a plant or plant component comprising the isolated nucleic acid sequence, a plant produced from a germinated seed of the plant, reducing the expression or enzymatic activity of a constitutive, or an ethylene induced or senescence induced tobacco polypeptide in a plant cell, and increasing the expression or enzymatic activity of a constitutive, or an ethylene or senescence induced tobacco polypeptide in a plant cell. The phenotypic trait comprises disease resistance, high yield, high grade index, curability, curing quality, mechanical harvestability, holding ability, leaf quality, height, maturation, stalk size, or leaf number per plant. The breeding method for producing a tobacco plant having decreased expression of a nicotine demethylase gene is useful developing desirable (non-genetically engineered) germplasm. The plant is useful in producing (smokeless) tobacco products. The tobacco product is a moist or dry snuff, a chewing tobacco, a cigarette product, a cigar product, a cigarillo, a pipe tobacco, or bidis. The p450 cDNAs were isolated using degenerate PCR primers designed against cytochrome p450 motifs. The present sequence is a cytochrome p450 cDNA of the invention.

Sequence 1538 BP; 464 A; 278 C; 334 G; 462 T; 0 U; 0 Other;

Alignment Scores:

Alignment scores:		
Pred. No.:	4.5e-102	Length: 1538
Score:	1164.00	Matches: 226
Percent Similarity:	63.4%	Conservative: 88
Best Local Similarity:	45.7%	Mismatches: 165

Query Match:	45.2%	Indels:	16
DB:	14	Gaps:	5
US-10-759-813-2 (1-500) x ABE06775 (1-1538)			
Qy	12	LeuIeSerPheLeuLeuValValLeuValValMetArgLeuTrpLysGln	31
Db	48	TTGGTTCCATTTTCCTATTTCTATCTTTCTTTTGTAAAGAAATGGAAGAACTCG	107
Qy	32	AsnPro-----ProProGlyProTrpLysPheProIleIleGlyAsnLeu	46
Db	108	AATAGCCAAAGAAAAATTGCCACAGCTCATGGAAATACCAATACTAGGAAGTATG	167
Qy	47	ProHisLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle	66
Db	168	CTTCATATGGTT-----GGTGGACTACCAACCACATGCTCTTAGAGATTTAGCCAAAAA	221
Qy	67	TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSerAla	86
Db	222	TATGGACGGCTTATGCACCTTCANTAGGTGAAGTTTCTGCAGTTGTGTTCTCTCT	281
Qy	87	GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal	106
Db	282	GATATGGCAAAAGAAAGTACTAAAACTCATGACATCGCTTTTCGGCTCTAGGCGCTT	341
Qy	107	LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp	126
Db	342	TTGGCCCGGAGATGTCTGTGTACAATAGGCTCATCTTGGCTTTTGGCCCTATGGCGAT	401
Qy	127	HisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysValGln	146
Db	402	TATTGGAGACAATCGTAAATATGTCTTGGAGTGTCTAGTGCCACAGAAATGTCGG	461
Qy	147	SerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLys	166
Db	462	ACATATAGCTCTATTAGCGCGATGAAGTCTTCGTCTCTCTTAATTTATCCGCTCATCT	521
Qy	167	AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleIleSerIleMetIle	186
Db	522	TCTGGTGAGCCTGTTAATATTACGGAAGGATCTTTTGTTCACAAGCTCCATGACATGT	581
Qy	187	ArgThrSerValGlyAsnCys--LysGlnLysGluArgLeuLeuSerValAlaAspAla	205
Db	582	AGATCAGCGTTTGGCGAAGTATTCAAGGACGACACAAATTTATACAACATAAAGAA	641
Qy	206	ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu	225
Db	642	GTTTATCTCTTAGCAGGAGGGTTTGTATGTGGCTGACATATTTCCTTCATACAAGCTCT	701
Qy	226	HisTyrIleIleGlyAlaGluSerLysProArgLeuHisGlnGluIleAspAspIle	245
Db	702	CATGTGCTCAGTGGAAATGAAGGTAAATATGAATGCACACCATAGGTAGTAGTCTATT	761
Qy	246	LeuGluGluIleLeuAsnGluHis-----LysAlaAsnLysPro	258
Db	762	GTTGAGAAATGTCATCAACGAGCACAGAAAAATCTTGCATTTGGGAAAACTAATGGAGCG	821
Qy	259	PheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValPro	278
Db	822	TTAGGAGGTGAAGATTTAAATTGATGTCTCTTAAAACTTATGAATGATGGAGCGCTTCAA	881
Qy	279	ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer	298
Db	882	TTTCCTATCACAACGACCACTCAAGCTATAATCTTTTGACATGTTTGTCTGTGGAACA	941
Qy	299	GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu	318
Db	942	GAGACTTCATCGTCAACAAATTTGTCTGGCTATGTTGGAAATGGTGGAAAAATCCAACGTGA	1001
Qy	319	LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu	338
Db	1002	TTTCGGAAGCTCAAGCAGAAGTAGAGATGCATTTTAGAGAAAAAGAAACTTTTGTATGAA	1061

Qy	398	ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr	417
Db	1242	TATTGGGATGATGCAGAAACATTTTAAGCCACAGAGATTGTCAAGAGTGTCTTAAGGATT	1301
Qy	418	LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIle	437
Db	1302	GTTGGTAATAATTTTGAATATCTTCCATTTTGGTGGAGGAGAGATTTGTCCAGGATTT	1361
Qy	438	ThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrp	457
		::: :::	
Db	1362	TCGTTTGGTTTTAGCTAATGCTTATTGTGCATTTGGCTCAATTACTTTATCACTTTGATTGG	1421
Qy	458	GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAla	477
Db	1422	GAATCCCCACTGGATCAACACCAAGCGCACTTGGACTTGACTGAGTTGGTTGGAGTAAC	1481
Qy	478	LeuArgLysIleAspLeuLysLeuIleProIleProTyrGln	492
		::: :::	
Db	1482	GCCGCTAGAAAAAGTGACCTTTACTTGGTCCGACTCCTTATCAA	1526

SQ	Sequence	1673 BP; 494 A; 309 C; 356 G; 514 T; 0 U; 0 Other;
 Alignment Scores:		
Pred. No.:	1..41e-100	Length: 1673
Score:	1149.00	Matches: 222
Percent Similarity:	63.2%	Conservative: 91
Best Local Similarity:	44.8%	Mismatches: 166
Query Match:	44.6%	Indels: 16
DB:	12	Gaps: 5
 US-10-759-813-2 (1-500) x ADN10646 (1-1673)		
Qy	12	LeulleSerPheLeuLeuValVallleLeuValValMetArgLeuTrpIysLysGln 31
Db	54	TGTGGTTCCATTTTCCTATTCTAFTCTTCTCTTTTTTGTTAAAGTATCGGAACACTCC 113
Qy	32	AsnPro-----ProGlyProTrpIysPheProfileileGlyAsnLeu 46
Db	114	AATAGCCAAAGCAAAAGTTGCCACCAGTCCATCGAACATACCAATACTAGGAAGTAG 173
Qy	47	ProHisLeuLeuThrSerAspLeuGlyHisGluArGPheArgAlaLeuAlaGlnIle 66
Db	174	CTTCATATGGTT-----GGTGCAGCTACCAACCATGCTTAGAGATTAGCCAAAAA 227
Qy	67	TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValIalleSerSerAla 86
Db	228	TATGGACCACCTATATGCACCTTCAATATAGTGAAGTTTCTGCGTTGTGTACTTCTCCT 287
Qy	87	GluAlaAlatysGluValMetIysThrGlnAlaAspAlaPheAlaGlnArgPrIleVal 106
Db	288	GATACGGCAAAAGAAGTATTAATAACTCATGACATCGCTTTTTCGCTTAGGCCCTAGCCCT 347
Qy	107	LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
Db	348	TTGCCCCGGAGATGTCTGTACATAGGTCTGATCTAGCCCTTTTGGCCCATGTTTCGG 407
Qy	127	HisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaIylsLysValGln 146
Db	408	TATTGGAGACAATATTAATGTACGGAAAGGATCTTTTGTTCACAAAGTCTCAGTGCCAAGAAATGTTTCGG 467
Qy	147	SerSerArgLeuIleArgGluGluMetGluaspAlaIleThrPheLeuArgSerIys 166
Db	468	ACATTTAGCTCTATTAGCGGAAATGAAGTTCTTCGCTCATTAATTTTATCCGGTCACT 527
Qy	167	AlaGlySerProValAsnIleThrLysIleIleTyrglyIleIleIleSerIleMetIle 186
Db	528	TCTGGTGNACCTATTAAATGTACGGAAAGGATCTTTTGTTCACAAAGTCTCCATGACATGT 587
Qy	187	ArgThrSerValGlyAsnCys---LysGlnLysGluArgLeuLeuSerValAlaAspAla 205
Db	588	AGATCAGCGTTTGGGCAAGTGTCAAAGAGCAAGCAAAAATTTATACAACTAATTAAGAA 647
Qy	206	ValAsnGluAlaAlaThrSerPheGlyThrAlaaspAlaPheProThrTrpIysLeuLeu 225
Db	648	GTGATACTCTTAGCAGGAGGGTGTGATGTGCTGCATATATTCCTTCTACTGAAGTTTCTT 707
Qy	226	HistyrIleIleGlyAlaGluSerIysProargLeuHisGlnGluIleAspAspIle 245
Db	708	CATGTGCTCAGTGAATGAAGGTGAAGATTATGANTGACACCATAGGTAGTAGTCCCAT 767
Qy	246	LeuGluGluIleLeuAsnGluHis-----LysAlaAsnLysPro 258
Db	768	GTTGAGAATGTCATCAATGACACCAAGAAAAATCTTGCAATTGGGAAAACTAATGGAGCG 827
Qy	259	PheGluAlaAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValPro 278
Db	828	TTAGGAGGTGAAGATTTAAATATGATGTCTTCTTAAGACTTATGAATGATGGAGCCTTCAA 887
Qy	279	ValProValThrAsnGluSerIleIyleValaSerValLeuGlnMetPheThrAlaGlySer 298
Db	888	TTTCCTATCACCAACGACACATCAAGCCATAATTTTGTGACATGTGTCGCCGGGACA 947
Qy	299	GluThrThrSerLysAlaThrGluTrpValIMetAlaGluLeuMetLysAsnProThrGlu 318

```
Db 948 GAGACTTCATCTCAACAAATTGTGGGCTATGGTAGAAATGGTGAATAATCCAGCCGTA 1007
Qy 319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
Db 1008 TTCGGCAAGCTCAGCAGAGTAGAGAGCAATTTAGAGGAAAGAAACATTTTCGATGAA 1067
Qy 339 SerArgPheHisAspLysPhePheLysLeuValLysGluThrLeuArgLeuHis 358
Db 1068 AATGATGTGGAGGAGTAAACTACCTAAAGTTAGTAAATAAAGAAACCTCTAAGACTTCAT 1127
Qy 359 ProProVal---ValLeuLeuProArgGluCysArgGluThrThrArgIleAspGlyTyr 377
Db 1128 CCACGGGTTCACATTTTGGTCCCAAGAGAAATGTAGGGAAGAGACAAATATAAACGGCTAC 1187
Qy 378 GluIleHisProAsnThrArgIleValValAenAlaTrpAlaIleGlyArgAspProAsn 397
Db 1188 ACTATTCCTGTAAAGACCAAGCTCATGTTAATGTTTGGCTTTGGGAGAGATCCAAA 1247
Qy 398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr 417
Db 1248 TATTGGAATGACGCAAGAACTTTTATGTCAGAGAGATTTGAGCAGTGTCTTAAGGATTTT 1307
Qy 418 LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIle 437
Db 1308 GTTGTGATAATTTTGAATATCTTCATTTGTGTGGCGGAAGAGGATTTGTCTGGGATT 1367
Qy 438 ThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuTyrHisPheAsnTrp 457
Db 1368 TCGTTTGGCTTAGCTTAATGCTTATTGGCAATGGCTCAATTACTATATCACTTCGATTGG 1427
Qy 458 GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyAla 477
Db 1428 AAATCCCTCGTGGAAATCGAACCAAGCGACTTGGACTTGACTGAGTTGGTTGGAGTAAC 1487
Qy 478 LeuArgLysLysIleAspLeuLysLeuIleProIleProTyrGln 492
Db 1488 GCCGCTAGAAAAAGTGACCTTTACTTGGTTGGCACTCCTCTTATCAA 1532
RESULT 8
ADN10626
ID ADN10626 standard; cDNA; 1673 BP.
XX
AC ADN10626;
XX
DT 15-JUL-2004 (first entry)
XX
DE Nicotiana p450 enzyme encoding cDNA SEQ ID NO:211.
XX
KW plant; p450; enzyme; Nicotiana; plant growth regulant; gene therapy;
KW phenotype; gene; ss.
XX
OS Nicotiana sp.
XX
PN WO2004035745-A2.
XX
PD 29-APR-2004.
XX
PF 16-OCT-2003; 2003WO-US032722.
XX
PR 16-OCT-2002; 2002US-0418933P.
PR 08-JUL-2003; 2003US-0485368P.
PR 18-SEP-2003; 2003US-0503989P.
XX
PA (USSW-) US SMOKELESS TOBACCO CO.
XX
PI Xu D;
XX
XX WPI; 2004-348441/32.
DR P-PSDB; ADN10627.
XX
PT New isolated nucleic acid molecules and encoded cytochrome p450 enzymes
from Nicotiana plants, useful for altering plant phenotypes.
```

```
XX Claim 1; SEQ ID NO 211; 198pp; English.
PS
XX The present sequence represents a plant p450 enzyme nucleic acid molecule
isolated from Nicotiana. Also described: (i) an isolated p450 protein
from Nicotiana; (2) a transgenic plant comprising the nucleic acid
molecule described above; (3) a method of producing the transgenic plant,
comprising: (i) operably linking the above nucleic acid molecule with a
promoter functional in the plant to create a plant transformational
vector; (ii) transforming the plant with the vector; (iii) selecting a
plant cell transformed with the transformation vector; and (iv)
regenerating a transformation plant from the transformed plant cell; and
(4) a method of selecting a plant containing the above nucleic acid
molecule, where the plant is analysed for the presence of the above
nucleic acid sequences. The p450 sequences have plant growth regulant
activity, and can be used in gene therapy. Compositions and methods from
the present invention are useful for altering plant phenotypes.
XX Sequence 1673 BP; 494 A; 309 C; 356 G; 514 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.: 1.41e-100 Length: 1673
Score: 1149.00 Matches: 222
Percent Similarity: 63.2% Conservative: 91
Best Local Similarity: 44.8% Mismatches: 166
Query Match: 44.8% Indels: 16
DB: 12 Gaps: 5
```

US-10-759-813-2 (1-500) x ADN10626 (1-1673)

```
Qy 12 LeuIleSerPheLeuLeuValLeuValValMetArgLeuTrpLysGln 31
Db 54 TTGTTTTCCATTTTCCTATCTTCCTTTTGTGAAGGATATGAAGAATCC 113
Qy 32 AsnPro-----ProProGlyProTrpLysPheProIleIleGlyAsnLeu 46
Db 114 AATAGCCAAAGCAAAAGTTGCCACCAGTCCATGTGNAACCTACCAATACTAGGAAGTATG 173
Qy 47 ProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle 66
Db 174 CTTCATATGGTT-----GGTGGACTACCAACACCATGTCTTAGAGATTTAGCCAAAAA 227
Qy 67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValAlaValIleSerSerAla 86
Db 228 TATGACCACCTTATGCACCTTCAATAGGTGAAGTTTCTGCGTTGCGTTACTTCTCTCT 287
Qy 87 GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
Db 288 GATACGGCAAAAGAGATATTAAAAAATCATGACATCGCTTTTTCGCTCTAGGCCTAGCCTT 347
Qy 107 LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
Db 348 TTGGCCCGGAGATTGTCTGTACAAATAGGTCTGTATCTAGCCTTTTGGCCCTTATGCGCAC 407
Qy 127 HisTrpArgGlnMetLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln 146
Db 408 TATTGGAGACAAATCCGAAATATGCTCTCTGGAAGTGTCTCAGTCCCAAGAAATGTTCCG 467
Qy 147 SerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLys 166
Db 468 ACATTTAGCTCTATTAGCGGAATGAAGTTCTTCGTCTCATTAATTTATCCGGTCATCT 527
Qy 167 AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleIleSerIleMetIle 186
Db 528 TCTGGTGAACCTATTAAATGTTACGAAAGGATCTTTTGTTCACAAGCTCCATGATGT 587
Qy 187 ArgThrSerValGlyAsnCys---LysGlnLysGluArgLeuLeuSerValAlaAspAla 205
Db 588 AGATCAGCGTTTGGCAAGTGTTCAGAGAGCAAGCAAAATTTATACAACTAATTAAGAA 647
Qy 206 ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu 225
Db 648 GTGATACTCTTAGCAGGAGGGTTTGATGTGGCTGACATATTCCTTCCCTTCACTGAAGTTTCT 707
```



```
QY 147 SerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLys 166
DB 468 ACATTAGCTCTATTAGCGGAATGAAGTTCTCTCTCATTAATTTATCGGTCATCT 527
QY 167 AlaGlySerProValAsnThrLysLysIleIleTyrGlyIleIleSerIleMetIle 186
DB 528 TCTGGTGAACTTATTAAAGTTAGGAAGGATCTTTTGTTCACAAGTCCATGACATGT 587
QY 187 ArgThrSerValGlyAsnCys---LysGlnLysGluArgLeuLeuSerValAlaAspAla 205
DB 598 AGATCAGCGTTTGGCGAAGTGTCAAAGAGCAAGCAAAATTTATACAACTAATTAAAGAA 647
QY 206 ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu 225
DB 648 GTGATACTCTTAGCAGGAGGTTTGATGTGGCTGACATATTCCTCTCACTGAAGTTCTT 707
QY 226 HisTyrIleIleGlyAlaGlySerLysProArgGluLeuHisGlnGluIleAspAspIle 245
DB 708 CATGTGCTCAGTGAATGAGGTAAGATTATGAATGCACACCATAGGTAGTCCCAT 767
QY 246 LeuGluGluIleLeuAsnGluHis-----LysAlaAsnLysPro 258
DB 768 GTTGAGAATGTCATCAATGAGCACAGAAAAATCTTGCAATTTGGGAAAACTAATGAGCG 827
QY 259 PheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValPro 278
DB 828 TTAGGAGGTGAAGATTAAATCATGTTCTTCAAGACTTATGAATGATGGAGCGCTTCAA 887
QY 279 ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer 298
DB 888 TTTCTATCACCACGACACATCAAGCCATAATTTTGCATCATGTTGCTCGCGGACA 947
QY 299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
DB 948 GAGACTTCATCGTCAACAAATTTGTGGCTATGTTAGTAAATGTTGAAAAATCCAGCGTA 1007
QY 319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
DB 1008 TTCGCGAAAGCTCAAGCAAGTAAGTAAGAGAGCATTTAGAGGAAAAAGAACTTTTCATGAA 1067
QY 339 SerArgPheHisAspLeuLysPheLysLeuValLysGluThrLeuArgLeuHis 358
DB 1068 ATGATGTGGAGGAGCTAACTACCTAAAGTTAGTAATAAAGAAACCTAAGACTTCAT 1127
QY 359 ProProVal---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyr 377
DB 1128 CCACCGGTTCCACTTTTGCTCCCAAGAGAAATGTAGGGAAGACAAATATAAACGGCTAC 1187
QY 378 GluIleHisProAsnThrArgIleValValAsnAlaTtpAlaIleGlyArgAspProAsn 397
DB 1188 ACTATTCTCTGTAAGACCAAGTCAATGTTTGGGCTTTGGGGAAGAGATCCAAA 1247
QY 398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr 417
DB 1248 TATTGGAATGAGCGACAAACTTTTATGCCAGAGAGATTTGACGAGTCTCTAAGGATTT 1307
QY 418 LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIle 437
DB 1308 GTTGGTAATAATTTTGAATATCTTCCATTTGTTGGCGGAAGGAGGATTTGCTGGGATT 1367
QY 438 ThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuTyrHisPheAsnTyr 457
DB 1368 TCGTTTGGCTTAGCTAATGCTTATTGTCATTTGGCTCAATTAATCTATATCACTTCGATTGG 1427
QY 458 GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAla 477
DB 1428 AAATCTCCCTGGGATCGAAACCAAGCGACTTTGGACTTGACTGAGTTGGTTGGAGTAAC 1487
QY 478 LeuArgLysLysIleAspLeuLysLeuIleProIleProTyrGln 492
DB 1488 GCCGCTAGAAAAAGTGACCTTTACTTTGGTTGGGACTCTCTTATCAA 1532
```

RESULT 10

```
ADZ65314
ID ADZ65314 standard; cDNA; 1673 BP.
XX AC ADZ65314;
XX 14-JUL-2005 (first entry)
XX Tobacco full length cDNA for cytochrome p450 clone D209-AA12.
XX Cytochrome p450; ss; secondary metabolite; ethylene; plant; senescence;
XX nornicotine; transgenic plant; gene.
XX Nicotiana tabacum.
XX WO2005038033-A2.
XX 28-APR-2005.
XX 15-OCT-2004; 2004WO-US034065.
XX 16-OCT-2003; 2003US-00686947.
XX 29-APR-2004; 2004US-0566235P.
XX 03-SEP-2004; 2004US-00934944.
XX (USSM-) US SMOKELESS TOBACCO CO.
XX Xu D;
XX WPI; 2005-315717/32.
XX P-PSDB; ADZ65315.
XX New nucleic acid molecule encoding cytochrome p450 enzymes in Nicotiana,
XX useful in developing tobacco plants with altered phenotypes.
XX Disclosure; SEQ ID NO 211; 226pp; English.
XX The invention relates to an isolated nucleic acid molecule (I) from
XX Nicotiana, where the nucleic acid molecule comprising any of the 59
XX nucleic acid sequences of SEQ ID NOS: 299-357 (NOTE: The claims refer to
XX SEQ ID NOS 299-357 as nucleic acids but these sequences (apart from SEQ
XX ID NO 356) are all proteins and appear as ADZ65402-ADZ65460. The nucleic
XX acids of the invention encode cytochrome p450 enzymes whose expression is
XX induced by ethylene and/or plant senescence. Also included are a
XX transgenic plant comprising (I), a method of producing a transgenic
XX plant, a method of selecting a plant containing a nucleic acid molecule
XX (where the plant is analyzed for the presence of nucleic acid sequence of
XX ADZ65402-
XX SQ Sequence 1673 BP; 494 A; 309 C; 356 G; 514 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.41e-100 Length: 1673
Score: 1149.00 Matches: 222
Percent Similarity: 63.2% Conservative: 91
Best Local Similarity: 44.8% Mismatches: 166
Query Match: 44.6% Indels: 16
DB: 14 Gaps: 5

US-10-759-813-2 (1-500) x ADZ65314 (1-1673)
QY 12 LeuIleSerPheLeuLeuValLeuIleLeuValValMetArgLeuTrpLysGln 31
DB 54 TTGGTTTCATTTTCTATTTCTATCTTCTCTTCTTTTGTAAAGATATGGAAGAACTCC 113
QY 32 AsnPro-----ProGlyProTrpLysPheProIleLeuGlyAsnLeu 46
DB 114 AATGCCAAAGCAAAAGTTGCCACGAGTCCATGCAAACTACCAATACTAGGAAGATG 173
QY 47 ProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle 66
DB 174 CTTTCATATGGTT-----GGTGGAGTACCACCACTGCTCTTAGAGATTACCAAAAAA 227
QY 67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValIleSerAla 86
```


XX SQ Sequence 1673 BP; 494 A; 309 C; 356 G; 514 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.4le-100 Length: 1673
Score: 1149.00 Matches: 222
Percent Similarity: 63.2% Conservative: 91
Best Local Similarity: 44.8% Mismatches: 166
Query Match: 44.6% Indels: 16
DB: 14 Gaps: 5

US-10-759-813-2 (1-500) x ADZ65861 (1-1673)

Qy		12	L <u>eutlSerPheLeu</u> LeuVal <u>leulleValMetArgLeuTrpLysLysGln</u> 31
Db		54	TGGTTCCTATTTCCATAATTCTAATCCTTTTGTTAAAGGATATGAAGAACCTCC 113
Qy		32	A <u>snpPro-----</u> - <u>ProProGlyPrOTrpLysPhoIrlleileGlyAsnLeu</u> 46
Db		114	AATAGCCAAGCAAAAAGTTGCCACCAGGTCCATGGAAACTACAATACTAGGAAGTAGT 173
Qy		47	P <u>rOHisLeuLeuThrSeraspLeuGlyHISgluaRgPHeArGalauLaGlinlle</u> 66
Db		174	CCTCATATGGTT-----GGTGGACTACCACACCATGTCTTAGAGATTATGCCAAAAA 227
Qy		67	TyrGlyP <u>rOvalMetSerLeugInlllegLnValSerAlavaiVailteSerSera</u> 86
Db		228	TATGGACCACCTTATGCACCTTCANATAGGTGAAGTTTTCHCGGTTGTGGTTACTTCTCT 287
Qy		87	GluAla <u>laLyLSguLValMetlysThrGlnlaAspAlaPheaGlnaRrgProfile</u> 106
Db		298	GATACGGCAAAGAGATTAATAAAACTCATGACATCGCTTTTGGCTAGGCCCTAGCCCTT 347
Qy		107	L <u>euaspa</u> laGlnlle <u>vaiPheTYrsenARgLysApvValLeuPheNalaseryTyRGlyAsp</u> 126
Db		348	TTGGCCCCGGAGATTGTCTGTACAAATAGTCTGATCTAGCTTTTGGCCCCTATGGCGCAC 407
Qy		127	HISTPA <u>RgInMetylsylslertprilleLeuGluPheLeuSerAlalyLysValGln</u> 146
Db		408	TATTGGAGACAAATCGGTAAATATGTCTTGGAAAGTCTCAGTGCCCAAGAAATGTTCCGG 467
Qy		147	SerSerArgLeu <u>IeARGGLuGluMetGluUspAlallethrPheLeuarGserLys</u> 166
Db		468	ACATTTAGCTCTATTAGCGCGAATGAAGTCTTCGTCTCATTAATTTTATCCGGTCATCT 527
Qy		167	Alagily <u>SerProvaAsnlleThrLyslilleITryGlylleileSerilleMetlle</u> 186
Db		528	TCTGTGTGAACCTATTAAATGTTACGGAAGGATCTTTTGTTCACAAGCTCCATGACATGT 587
Qy		187	Arg <thr< th="">SerValigLYasnCys---LysGinLysGUarGLeuLeuSerValala<u>AspAla</u> 205</thr<>
Db		598	AGATCAGCGTTTGGGCAAGTGTCAAAGAGCAAGACAAAATTTATACAACATTAATAAGAA 647
Qy		206	VallanGlu <u>aia</u> la <u>ThrSerpheGlyThrAlaaSpAlapheProthrrtpLysLeuLeu</u> 225
Db		648	GTGAPACTCTTAGCAGGAGGGTTGATGGCTGACATATTCCTTCATGAAGTTTCTT 707
Qy		226	HistYrIlleillegLYalaGluSerLysPRoarGLeuHiHGlnGluille <u>ASPaspille</u> 245
Db		708	CATGTCTCAGTGGGAATGNAGSGGTAAATATGAATGCACACATAAAGGTAGTAGCCATT 767
Qy		246	LeuGluGluille <u>LeuasngLuHis-----</u> --LysalahenLysPro 258
Db		768	ETTGAGAATGTCATCAATGAGCACAGAANAATCTTTGCCAATTTGGGAAAACTAATGGAGCG 827
Qy		259	pHeGluAla <u>ASpAsnLeuMetAppValLeuLeuAsnLeuGlnLysAsnGlyAsnValPro</u> 278
Db		828	TTAGGAGGTGAAGATTAAATGATGTTCTTCAAGACTTATGAATGATGGAGGCCCTCAA 887
Qy		279	ValProv <u>alThrAsnGlnSerilleysAlaSerValLeuLeuGlnMetPheThrAlaGlySer</u> 298
Db		888	TTTCTCTATCACCAACGACAACATCAAGGCCATAAATTTTGGACATGTTTGTCTGCCGGGACA 947

of nornicotine.

Disclosure; SEQ ID NO 231; 203pp; English.

The invention relates to an isolated nucleic acid molecule from Nicotiana, encoding a protein. The invention also relates to a transgenic plant comprising the nucleic acid molecule, a method of producing a transgenic plant comprising operably linking the nucleic acid molecule with a promoter functional in the plant to create a plant transformational vector, transforming the plant with the plant transformational vector, selecting a plant cell transformed with the transformation vector and regenerating a transformation plant from the transformed plant cell, a method of selecting a plant containing a nucleic acid molecule, a method of increasing or decreasing nicotine levels in a plant by operably linking the nucleic acid molecule with a promoter functional in the plant, a tobacco product having reduced amounts of nicotine levels, the tobacco product comprising tobacco from the plant, a tobacco leaf having reduced amounts of nicotine levels and a method of isolating a gene from a plant using the isolated nucleic acid. In producing a transgenic plant, the plant has reduced levels of nicotine. The tobacco product is selected from cigarettes, cigars, pipe tobacco, snuff, chewing tobacco, products blended with the tobacco product and their mixtures. The nucleic acid molecule is useful for altering plant phenotypes, thus producing a transgenic plant having reduced levels of nicotine. This sequence represents cDNA encoding a tobacco cytochrome P450 enzyme of the invention.

Sequence 1673 BP; 494 A; 309 C; 356 G; 514 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.41e-100	Length:	1673
Score:	1149.00	Matches:	222
Percent Similarity:	63.2%	Conservative:	91
Best Local Similarity:	44.8%	Mismatches:	166
Query Match:	14.6%	Indels:	16
DB:	14	Gaps:	5

US-10-759-813-2 (1-500) x ADZ65881 (1-1673)

Qy	12	LeuIleSerPheLeuLeuValLeuLeuLeuValValMetArgLeuTrpLysGln	31
Db	54	TTGGTTCCCATTTTCCTATCTTCTCTCTTTTAAAGTATGGAAGAACTCC	113
Qy	32	AsnPro-----ProGrgLYP-oTrpLYsPheProIleIleGlyAsnLeu	46
Db	114	AATAGCCAAACCAAAAGTTGCCACGATCATGGAAACTACCAATACTAGGAAGTATG	173
Qy	47	ProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle	66
Db	174	CTTCATATGGTT-----GGTGGACTACCAACCACATGCTCTTAGAGATTAGCCAAAAA	227
Qy	67	TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValIleSerSerAla	86
Db	228	TATGGACCATCTATAGCCTTCAATTAGGTGAAGTTTCTGGGTTGTGTACTTCTCTCT	287
Qy	87	GluAlaIatLysGluValMetLysThrGlnAlaaspAlaPheAlaGlnArgProIleVal	106
Db	288	GATACGGCAAAAGAAGTATTAAAACTCATGACATCGCTTTTGCGCTCTAGGCCTAGCCCT	347
Qy	107	LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp	126
Db	348	TTGGCCCCGGAGATTGCTGTTTCAATAAGGTCGATCTAGCCTTTTGCCCCCTATGGCAG	407
Qy	127	HisTrpArgGlnMetLysLysIleLeuTrpIleLeuGluPheLeuSerAlaLysLysValGln	146
Db	408	TATTGGAGACAAATCGGTAAATATGTCTTTGGAAGTGTCTCAGTGCCCAAGAAATTTTCGG	467
Qy	147	SerSerArgLeuIleArgGluGluMetGluaspAlaIleThrPheLeuArgSerLys	166
Db	468	ACATTTAGTCTATTAGCGGAATGAAGTCTTCGTCTCATTAATTTTATCCGGTCATCT	527
Qy	167	AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleIleSerIleMetIle	186

RESULT 13

RECEIVED 12
AEE06853
ID AEE06853 standard: CDNA: 1673 BP.

AA AEE06853:

AC
XX
DT

QY	299	GlUThrThrSerIysAlaThrGlUThrValMetAlaGlULeuMetIysAsnProThrGlu	318
		:::	
Db	948	GAGACTTCATCGTCAACAATTTGTGTGGCTATGTTAGAAATGGTAAAAATCCAGCGGTA	1007
		:::	
QY	319	LeuArgIysAlaGlnGlUValArgGlnValPheGlyGluMetGlyIysValAspGlu	338
		:::	
Db	1008	TTCCGGAAGACTCAAGCAGAAGTAAGAGACATTTTAGAGAAAGAAACTTTTCGGATGAA	1067
		:::	
QY	339	SerArgPheHisAspLeuLysPhePheLysLeuValValIysGluThrLeuArgLeuHis	358
		::: :::	
Db	1068	AATGATGTGGAGAGCTAAACTCTAAAGTTAGTAATAAAGAAACTCTAAAGACTTCAT	1127
		:::	
QY	359	ProProVal---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyr	377
		:::	
Db	1128	CCACCGGTTCCACTTTTGTCTCCCAAGAGAGATGTAGGGAAGAGACAAATATAAACGGCTAC	1187
		:::	
QY	378	GluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsn	397
		:::	
Db	1188	ACTATTCTCTAAGACCAAAGTCATGGTTAAATGTTTGGGCTTTGGGAAGAGATCCAAAA	1247
		:::	
QY	398	ThrTpsSerGluProGlyIysPheAsnProGluArgPheIysAspCysAlaIleAspTyr	417
		:::	
Db	1248	TATTTGGAATGACGCAGCAAACTTTTATGCCAGAGAGATTTTGACGAGTCTCTAAGGATTTT	1307
		:::	
QY	418	LysGlyThrThrPheGluLeuValProPheGlyValaGlyIysArgIleCysProGlyIle	437
		:::	
Db	1308	GTTGGTAATAATTTTGAATATCTTCCTATTGGTGGCGAAGAGGATTTTCTCTGGGANT	1367
		:::	
QY	438	ThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTyr	457
		::: :::	
Db	1368	TCGTTTGGCTTAGCTAATAGCTTATTGCCAATGGCTCAATTACTATATCATCTTCGATTGG	1427
		:::	
QY	458	GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAla	477
		:::	
Db	1428	AAACTCCCTCGTGGAAATCGAACCAACGACCTTGGACTTGACTGAGTTGGTTGGAGTAAC	1487
		:::	
QY	478	LeuArgIysIysIleAspLeuLysLeuIleProIleProTyrGln	492
		:::	
Db	1488	CGCGCTAGAAAAAGTGCCTTTACTTTGGTTGGCACTCCTTATCAA	1532

RESULT 14

RECEIVED 11/1/94
AEE06833
ID AEE06833 standard; cDNA; 1673 BP.

AA
AC AEE06833:

26-JAN-2006 (first entry)

XX
DE
Tobacco p450 cDNA SEQ ID NO: 220.

XX Plant breeding; plant; ss; coding sequence; cytochrome p450;
KW transgenic plant; senescence.
KW

XX
QS
Nicotiana tabacum

XX PN WO2005111217-A2

XX
PD
24-NOV-2005XX
PF 27-APR-2005: 2005WQ-TIS014803XX
PR 29-APR-2004 2004US-0566235P

PR 03-SEP-2004; 2004US=00934944.
PR 03-SEP-2004: 2004US=0607357P

PR 17-SEP-2004; 2004US-00943507.
PR 15-OCT-2004; 2004WO-115034065

PR 25--JAN-2005: 2005IIS-0646764P
PR 15-OCT-2004; 2004WO-US034218.

PR 24-MAR-2005; 2005US-066509/P.
PR 24-MAR-2005. 2005US-066545/P.

XX
PA (USSM-) IIS SMOKELESS TOBACCO[illegible]

Xu D, Nielsen MT;
WPI; 2005-786788/80.
P-PSDB; AEE06834.

Producing a tobacco plant having decreased expression of a nicotine demethylase gene comprises crossing a first tobacco plant with a second tobacco plant and germinating the collected seed of an F1 progeny plant.

Claim 32: SEQ ID NO 220; 641pp: English:

The invention relates to a breeding method for producing a tobacco plant with reduced expression of a nicotine demethylase gene comprising crossing a first tobacco plant with variant nicotine demethylase gene expression with a second tobacco plant with at least one phenotypic trait to produce an F1 progeny plant, the seed of which is collected and germinated to produce a tobacco plant having reduced expression of a nicotine demethylase gene. Also included are breeding a nicotine demethylase deficiency trait into a tobacco plant, producing a tobacco seed, developing a tobacco plant in a tobacco breeding program, a tissue culture of regenerable tobacco cells obtained from the tobacco plant of the invention, producing a tobacco product, a breeding method for producing a tobacco plant having a modified attribute, a method of breeding an attribute into a tobacco plant, a tobacco plant or its components produced by the method of breeding a nicotine demethylase deficiency trait into a tobacco plant, producing tobacco seed, producing a tobacco plant having a modified attribute or developing a tobacco plant in a tobacco breeding program, an isolated genetic marker comprising a nucleic acid sequence that is substantially identical to a nucleic acid sequence given in the specification (the nucleic acids comprise isolated cytochrome p450 cDNAs), an expression vector comprising the nucleic acid sequence, a plant or plant component comprising the isolated nucleic acid sequence, a plant produced from a germinated seed of the plant, reducing the expression or enzymatic activity of a constitutive, or an ethylene induced or senescence induced tobacco polypeptide in a plant cell, and increasing the expression or enzymatic activity of a constitutive, or an ethylene or senescence induced tobacco polypeptide in a plant cell. The phenotypic trait comprises disease resistance, high yield, high grade index, curability, curing quality, mechanical harvestability, holding ability, leaf quality, height, maturation, stalk size, or leaf number per plant. The breeding method for producing a tobacco plant having decreased expression of a nicotine demethylase gene is useful developing desirable (non-genetically engineered) germplasm. The plant is useful in producing (smokeless) tobacco products. The tobacco product is a moist or dry snuff, a chewing tobacco, a cigarette product, a cigar product, a cigarillo, a pipe tobacco, or bids. The p450 cDNAs were isolated using degenerate PCR primers designed against cytochrome p450 motifs. The present sequence is a cytochrome p450 cDNA of the invention.

Sequence 1673 BP: 494 A: 309 C: 356 G: 514 T: 0 U: 0 Other: 0

Alignment Scores:

Alignment Scores:					
Pred. No.:	1,419-100	Length:	1673		
Score:	1149.00	Matches:	222		
Percent Similarity:	63.2%	Conservative:	91		
Best Local Similarity:	44.8%	Mismatches:	166		
Query Match:	44.6%	Indels:	16		
DB:	14	Gaps:	5		

US-10-759-813-2 (1-500) x AEE06833 (1-1673)

[illegible]

```
Qy 67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValIleSerSerAla 86
Db 228 TATGACCACTTATGACCTTCAATAGTAGTGAAGTTCTGCGTGTGTGTTACTTCTCT 287
Qy 87 GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
Db 288 GATACGGCAAGAAGTAGTATTAATAACTCATGACATCGCTTTTGGCGCTAGGCCTT 347
Qy 107 LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
Db 348 TTGGCCCGGAGATTGCTGTACAAATAGTCTGATCTAGCTTTTGGCCCTTAGCCGAC 407
Qy 127 HisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln 146
Db 408 TATTGGAGACAATCGTAAATAATGTGCTTTGGAAGTCTCAGTGCACAAGAATGTCGG 467
Qy 147 SerSerArgLeuIleArgGluGluGluMetGluAspAlaIleThrPheLeuArgSerLys 166
Db 468 ACATTTAGCTCTATTAGCGGGAATGAAGTTCTTCGTCTCATTAATTTTATCGGTCTCT 527
Qy 167 AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleIleSerIleMetIle 186
Db 528 TCTGTGTGAACCTATTATGTTAGGAAGGATCTTTTGTTCACAAGCTCCATGACATGT 587
Qy 187 ArgThrSerValGlyAsnCys---LysGlnLysGluArgLeuSerValAlaAspAla 205
Db 588 AGATCAGCGTTTGGCAAGTGTCAAGAGCAAGCAAAATTTATACAACCTAATTAAGAA 647
Qy 206 ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu 225
Db 648 GTGATACTCTTAGCAGAGGGTTTGATGTGGCTGACATATTCCTTCACTGAAGTTCTT 707
Qy 226 HisTyrIleIleGlyAlaGlnSerLysProArgLeuHisGlnGluIleAspAspIle 245
Db 708 CATGTGCTCAGTGAATGAGGTAGATATTAATGATGCACACCATAGGTAGTATCCATT 767
Qy 246 LeuGluGluIleLeuAsnGluHis-----LysAlaAsnLysPro 258
Db 768 GTTGAGAATGTCATCAATGAGCACAAGAAAAATCTTGCAATGGGAAAACTAATGGAGC 827
Qy 259 PheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValPro 278
Db 828 TTAGAGGAGTGAAGATTAAATGATGTTCTTAAGACTTATGAATGATGGAGCGCTTCAA 887
Qy 279 ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer 298
Db 888 TTTCTCTATCACCACGACCAACATCAAGCCATAATTTTGACATGTTTGTCTCCCGGACA 947
Qy 299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
Db 948 GAGACTTCATCGTCAACAATTTGTGGGTATGGTAGAATGGTGAAAAATCCAGCCGTA 1007
Qy 319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
Db 1008 TTCGCGAAGACTCAAGCAGAGTAGAGAGAGCATTTAGAGGAAAAAGAACTTTTCATGAA 1067
Qy 339 SerArgPheHisAspLeuLysPheLysLeuValLysGluThrLeuArgLeuHis 358
Db 1068 AATGATGTGGAGAGCTAACTACCTAAAGTTAGTAATAAAGAAGAACTTAAGACTTCAT 1127
Qy 359 ProProVal---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyr 377
Db 1128 CCACCGGTTCCACTTTTGTCTCCCAAGAGAAATGTAGGGAAGACAAATATAAACGGCTAC 1187
Qy 378 GluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsn 397
Db 1188 ACTATTCTCTGTGAAGACCAAACTGATGTGTATGTTTGGGCTTTGGGAAGAGATCAAAA 1247
Qy 398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr 417
Db 1248 TATTGGAATGACGCAAGAACTTTTATGCCAGAGAGATTTGACAGCTGCTCTAAGGATTTT 1307
```

```
Qy 418 LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIle 437
Db 1308 GTTGGTAATAATTTTGAATATCTTCATTTGGTGGCGAAGAGGATTTGTCTCGGATT 1367
Qy 438 ThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrp 457
Db 1368 TCGTTTGGCTTAGCTAATGCTATTGTCATTTGGCTCAATTACTATATCACTTCGATTGG 1427
Qy 458 GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAla 477
Db 1428 AAACCTCCCTGCTGGAATCGAACCAAGCGACTTGGACTTGAAGTTGGTTGGAGTAACT 1487
Qy 478 LeuArgLysLysIleAspLeuLysLeuIleProIleProTyrGln 492
Db 1488 CCCGCTAGAAAAAGTGACCTTTACTTGTGCTGCACTCCTTATCA 1532
RESULT 15
AEF97226 standard; cDNA; 1673 BP.
XX AEF97226;
XX 20-APR-2006 (first entry)
XX Tobacco cytochrome P450 cDNA clone D209-AA12, SEQ ID NO:220.
XX Plant breeding; crop improvement; secondary metabolite; genetic marker;
XX plant; cytochrome P450; CYP; gene; ss.
XX Nicotiana tabacum.
XX US2006041949-A1.
XX 23-FEB-2006.
XX 27-APR-2005; 2005US-00116881.
XX 13-NOV-2001; 2001US-0337684P.
XX 11-JAN-2002; 2002US-0347444P.
XX 12-MAR-2002; 2002US-0363684P.
XX 16-OCT-2002; 2002US-0418933P.
XX 13-NOV-2002; 2002US-00293252.
XX 10-JAN-2003; 2003US-00340861.
XX 12-MAR-2003; 2003US-00387346.
XX 08-JUL-2003; 2003US-0485368P.
XX 18-SEP-2003; 2003US-0503989P.
XX 16-OCT-2003; 2003US-00686947.
XX 29-APR-2004; 2004US-0566235P.
XX 03-SEP-2004; 2004US-00934944.
XX 03-SEP-2004; 2004US-0607357P.
XX 17-SEP-2004; 2004US-00943507.
XX 15-OCT-2004; 2004WO-US034065.
XX 15-OCT-2004; 2004WO-US034218.
XX 25-JAN-2005; 2005US-0646764P.
XX 24-MAR-2005; 2005US-0665097P.
XX 24-MAR-2005; 2005US-0665451P.
XX 19-APR-2005; 2005US-00110062.
XX (USSM-) US SMOKELESS TOBACCO CO.
XX Xu D, Nielsen MT;
XX WPI; 2006-182895/19.
XX P-PSDB; AEF97227.
XX New breeding method, useful for producing a tobacco plant having
XX decreased expression of a nicotine demethylase gene comprises crossing
XX germinating F1 progeny seed to produce the tobacco plant.
XX Claim 63; SEQ ID NO 220; Sllpp; English.
XX The invention relates to a breeding method for producing a tobacco plant
XX having decreased expression of a nicotine demethylase gene. The method
CC
```

Qy	12	LeuIleSerPheLeuValLeuValLeuValMetArgLeuTrpLysGln	31
Db	54	TtGGTTCCATTTCTTCTATCTTTTCCTTTTAAAGTATGGAAGACTCC	113
Qy	32	AsnPro-----ProProGlyProTrpLysPheProIleLeuGlyAsnLeu	46
Db	114	AATAGCCCAAGCAAAGATTGCCACAGGTCTCATGAAACTACCATACTAGAAGTAGT	173
Qy	47	ProHisLeuLeuThr-SerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle	66
Db	174	CttCATATGGTt-----GGTGAGCTACCAACCATGCTCTTAGAGATTAGCCAAAAA	227
Qy	67	TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValIleSerSerAla	86
Db	228	TATGGACCACTTTATGCACCTTCATTTAGGTGAAGTTTTCTGCGGTTGTGTCTTCTCCT	287
Qy	87	GluAlaAlaLysGluValMetIysThrGlnAlaAspAlaPheAlaGlnArgProIleVal	106
Db	298	GATACGGCAAAAGAAGTATTAAAANAACtCATGACATCGCTTTTTCGTCTAGGCCCTAGCCIT	347
Qy	107	LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp	126

Qy 478 LeuArgLysLysIleAspLeuLysLeuIleProileProTyrGln 492
Db 1488 GCCGCTAGAAAAAGTGACCTTTACTTGGTTGGGACTCCTTATCAA 1532

Search completed: May 31, 2006, 06:42:14
Job time : 939 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 31, 2006, 06:42:32 ; Search time 1730 Seconds
(without alignments)
5327.010 Million cell updates/sec

Title: US-10-759-813-2
Perfect score: 2574
Sequence: 1 MBQKNLSPSILISFLVLI.....KIDKLPIPIYQVSLGNSIS 500

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n model -DEV=xlh
-Q=abss/ABSSWEB_spool/US10759813/runat_31052006_060313_2784/app_query.fasta_1
-DB=Published Applications NA Main -QFMT=fastap -SUFFIX=rnpbm -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abss05h
-USER=US10759813 @CGN_1_1_1675@runat_31052006_060313_2784 -NCPU=6 -ICPU=3
-NO_WMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA Main:
1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq:*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq:*
15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq:*
16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length DB ID	Description
1	2574	100.0	1733	3 US-09-909-566C-1 Sequence 1, Appli

2	2574	100.0	1733	8	US-10-759-813-1	Sequence 1, Appli
3	1164	45.2	1538	8	US-10-686-947-153	Sequence 153, App
4	1164	45.2	1538	10	US-10-943-507-153	Sequence 153, App
5	1164	45.2	1538	11	US-10-934-944-153	Sequence 153, App
6	1164	45.2	1538	16	US-11-116-881A-162	Sequence 211, App
7	1149	44.6	1673	8	US-10-686-947-211	Sequence 211, App
8	1149	44.6	1673	8	US-10-686-947-231	Sequence 231, App
9	1149	44.6	1673	10	US-10-943-507-208	Sequence 208, App
10	1149	44.6	1673	11	US-10-934-944-211	Sequence 211, App
11	1149	44.6	1673	11	US-10-934-944-231	Sequence 231, App
12	1149	44.6	1673	16	US-11-116-881A-220	Sequence 220, App
13	1149	44.6	1673	16	US-11-116-881A-240	Sequence 240, App
14	1148	44.6	1673	8	US-10-686-947-209	Sequence 209, App
15	1148	44.6	1673	10	US-10-943-507-206	Sequence 206, App
16	1148	44.6	1673	11	US-10-934-944-209	Sequence 209, App
17	1148	44.6	1673	16	US-11-116-881A-218	Sequence 218, App
18	1144	44.4	1857	8	US-10-686-947-195	Sequence 195, App
19	1144	44.4	1857	10	US-10-943-507-192	Sequence 192, App
20	1144	44.4	1857	11	US-10-934-944-195	Sequence 195, App
21	1144	44.4	1657	16	US-11-116-881A-204	Sequence 204, App
22	1142	44.4	1673	8	US-10-686-947-213	Sequence 213, App
23	1142	44.4	1673	10	US-10-943-507-210	Sequence 210, App
24	1142	44.4	1673	11	US-10-934-944-213	Sequence 213, App
25	1142	44.4	1673	16	US-11-116-881A-222	Sequence 222, App
26	1131	43.9	1610	8	US-10-686-947-197	Sequence 197, App
27	1131	43.9	1610	8	US-10-686-947-261	Sequence 261, App
28	1131	43.9	1610	10	US-10-943-507-194	Sequence 194, App
29	1131	43.9	1610	10	US-10-943-507-256	Sequence 256, App
30	1131	43.9	1610	11	US-10-934-944-197	Sequence 197, App
31	1131	43.9	1610	11	US-10-934-944-261	Sequence 261, App
32	1131	43.9	1610	16	US-11-116-881A-206	Sequence 206, App
33	1131	43.9	1610	16	US-11-116-881A-270	Sequence 270, App
34	1129.5	43.9	1576	8	US-10-686-947-149	Sequence 149, App
35	1129.5	43.9	1576	10	US-10-943-507-149	Sequence 149, App
36	1129.5	43.9	1576	11	US-10-934-944-149	Sequence 149, App
37	1129.5	43.9	1576	16	US-11-116-881A-158	Sequence 158, App
38	1128.5	43.8	1566	8	US-10-686-947-191	Sequence 191, App
39	1128.5	43.8	1566	10	US-10-943-507-188	Sequence 188, App
40	1128.5	43.8	1566	11	US-10-934-944-191	Sequence 191, App
41	1128.5	43.8	1566	16	US-11-116-881A-200	Sequence 200, App
42	1128.5	43.8	1581	8	US-10-686-947-151	Sequence 151, App
43	1128.5	43.8	1581	10	US-10-943-507-151	Sequence 151, App
44	1128.5	43.8	1581	11	US-10-934-944-151	Sequence 151, App
45	1128.5	43.8	1581	16	US-11-116-881A-160	Sequence 160, App

ALIGNMENTS

RESULT 1
US-09-909-566C-1
; Sequence 1, Application US/09909566C
; Publication No. US20030066103A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Edgar B
; TITLE OF INVENTION: A Cytochrome P450 enzyme associated with the synthesis of delta-
; FILE REFERENCE: BBI465 US NA
; CURRENT APPLICATION NUMBER: US/09/909,566C
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 60/219833
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1733
; TYPE: DNA
; ORGANISM: Euphorbia lagascae
US-09-909-566C-1

Alignment Scores:	6.66e-282	Length:	1733
Pred. No.:	2574.00	Matches:	500
Score:	100.0%	Conservative:	0


```
QY 41 ProIleGlyAsnLeuProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPhe 60
DB 134 CCTATCATAGTAATCTTCTCATTTATTACTACATCTTGATCTAGCCCATGAACGTTTT 193
QY 61 ArgAlaLeuAlaGlnIleTyrGlyProValMetSerLeuGlnIleGlyGlnValSerAla 80
DB 194 AGAGCCCTGGCTCAAAATTTATGGACCTGTTATGAGTCTTCAAAATGGCCCAAGTTTCAGCT 253
QY 81 ValValIleSerSerAlaGluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPhe 100
DB 254 GTTGTCATTTCTTCAGCTGAAGCAGCCAAAGAGGTTATGAAAACCTCAGGCTGATGCCCTTC 313
QY 101 AlaGlnArgProIleValLeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeu 120
DB 314 GCCCAACGCCCTATCGCTTGGACGCACAGATGTGTTTTATTAATCGGAAGAATGTCCTG 373
QY 121 PheAlaSerTyrGlyAspHisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeu 140
DB 374 TTTGCTTCATATGGAGATCACTGGAGGCAGATGAAGAAAAATTTGGATACTTGAATTTCTG 433
QY 141 SerAlaLysLysValGlnSerSerArgLeuIleArgGluGluMetGluAspAlaIle 160
DB 434 AGTGCCAAATAAAGTTCAATCTCTCAGGTTAATCCGAGAGGAAGAAATGGAGGATGCCATC 493
QY 161 ThrPheLeuArgSerLysAlaGlySerProValAsnIleThrLysIleIleTyrGlyIle 180
DB 494 ACATCTCTCGTTCGNAACCGGATCTCCGGTCAATATTACAAAGATCATTTATGGCAT 553
QY 181 IleIleSerIleMetIleArgThrSerValGlyAsnCysLysGlnLysGluArgLeuLeu 200
DB 554 ATAATTTCCATCATGATAAGAACATCCGTTGGTAAATTTGAAGCAAAAGAAAGATTTGCTG 613
QY 201 SerValAlaAspAlaValAsnGluAlaThrSerPheGlyThrAlaAspAlaPhePro 220
DB 614 AGTGTCGCATGCGCATCAATGAGGAGCGAGGTTTTGGCACCCGACGCTTTTCG 673
QY 221 ThrTrpLysLeuLeuHisTyrIleIleGlyAlaGluSerLysProArgArgLeuHisGln 240
DB 674 ACGTGGAAATTACTTCACATATATCATTTGCGAGCTGAGTCAAAACCCAGCGGTTTGCATCAG 733
QY 241 GluIleAspAspIleLeuGluGluLeuLeuAsnGluHisLysAlaAsnLysProPheGlu 260
DB 734 GAGATTGACGATATATCTGAAGAGATTCTTAATGAACAAAGCCCAATAAGCCTTTTGA 793
QY 261 AlaAspAsnLeuMetAspValLeuLeuAsnGlnLysAsnGlyAsnValProValPro 280
DB 794 GCGGATAACTTAATGGATGTTCTATTGAATCTTCAAAAAAATGGAAACGTTCCAGTGCCA 853
QY 281 ValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySerGluThr 300
DB 854 GTGACAAACGNAAGCATCAAGCATCCGTTTGGCAATGTTTACTGCGGGAGCGAACA 913
QY 301 ThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGluLeuArg 320
DB 914 ACTTCGAAAGCTACAGAATGGGTAAATGGCAGAGCTGATGAAAAATCCAACTGAACCTAAGA 973
QY 321 LysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGluSerArg 340
DB 974 AAAGCACAAAGAGAAGTTAGACAAGTATTGGTGAATGGGAAAAAGTTGATGAATCAAGA 1033
QY 341 PheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHisProPro 360
DB 1034 TTTTCATGTTGAATTTCTCAAGTAGTGGTTAAAGAAACTCTAAGATTACATCCTCG 1093
QY 361 ValValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyrGluLeHis 380
DB 1094 GTTGCTTGTATTCGAGGAGGTGTAGAGAAACAACACAGAAATGATGATGAATGAATTCAT 1153
QY 381 ProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsnThrTrpSer 400
DB 1154 CCGAACACTCGAATTTGTGTGAATGCTTGGGCGATAGGAAGAGATCCCTAATCTTGGTGC 1213
```

```
QY 401 GluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyrLysGlyThr 420
DB 1214 GAACCTGGAAAGTTTAAACCAGAGAGTTTAAAGATTGTGCAATTGATTATAAAGGACG 1273
QY 421 ThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIleThrSerAla 440
DB 1274 ACATTTGAACCTGGTACCATTTGGTGCAGGAAAAAAGAAATATGCTCGCATTTACTCAGCT 1333
QY 441 IleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrpGluLeuAla 460
DB 1334 ATTACCAATTTGAGTATGTCAATTATAAATCTATTATATCAATTTAATTCGGAACTGGCC 1393
QY 461 AspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAlaLeuArgLys 480
DB 1394 GATGGAATTTACACCTCAACACCTTGATATGACTGGAAGCTATTGGCGGTGCTCTCAGGAAA 1453
QY 481 LysIleAspLeuLysLeuIleProIleProTyrGlnValSerLeuGlySerAsnIleSer 500
DB 1454 AAAATAGATCTTAAAGTTGATTCTCTATTCCATATCAAGTTAGCTTAGCTCAAAATATTCT 1513
RESULT 3
US-10-686-947-153
; Sequence 153, Application US/10686947
; Publication No. US20040162420A1
; GENERAL INFORMATION:
; APPLICANT: Profigen Inc.
; TITLE OF INVENTION: Cloning of Cytochrome P450 Genes from Tobacco
; FILE REFERENCE: 79601
; CURRENT APPLICATION NUMBER: US/10/686,947
; CURRENT FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: US 10/387346
; PRIOR FILING DATE: 2003-03-12
; NUMBER OF SEQ ID NOS: 298
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 153
; LENGTH: 1538
; TYPE: DNA
; ORGANISM: NICOTIANATABACUM
US-10-686-947-153
Alignment Scores:
Pred. No.: 2,77e-121 Length: 1538
Score: 1164.00 Matches: 226
Percent Similarity: 63.4% Conservative: 88
Best Local Similarity: 45.7% Mismatches: 165
Query Match: 45.2% Indels: 16
DB: Gaps: 5
US-10-759-813-2 (1-500) x US-10-686-947-153 (1-1538)
QY 12 LeuIleSerPheLeuLeuValLeuIleLeuValValValMetArgLeuTrpLysGln 31
DB 48 TTGGTTTCCATTTCTTCTATTCTTTCTTTTCTTTTAAAGAAATGGAAGAACTCG 107
QY 32 AsnPro-----ProProGlyProTrpLysPheProIleIleGlyAsnLeu 46
DB 108 ATATGCCAAAGCAAAATAATTCGCCACCATGCTCCATGGAACCTACCAATACATAGGAAGTATG 167
QY 47 ProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle 66
DB 168 CTTTCATATGGTT-----GGTGGACTACCAACCATGCTCTTTAGAGATTTAGCCAAAAA 221
QY 67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValIleSerSerAla 86
DB 222 TAAGACCGCTTATGACCTTCATTTAGGTGAAGTTTCTGCAGTTGTGTCTTCTCTCCT 281
QY 87 GlnAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
DB 282 GATATGGCAAAAGAAAGTACTAAAAACTCATGACATCGCTTTTCGGTCTAGGCCCTAGCCCTT 341
QY 107 LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
DB 342 TTGGCCCGGAGATTGTCTGTACAAATAGTCTGATCTTGTGCGTTTGTGCCCTATGCGCAT 401
```

```
Qy 127 HistPrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln 146
Db 402 TATTGGAGACAAATCGTAAATATGTGCTTGGAGAGTCTCAGTGCACAAAGATGTCGG 461
Qy 147 SerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLys 166
Db 462 ACATATAGCTCTATTAGCGCATGAAGTCTTCCTGCTCTTAATTTATTCGGGTCACT 521
Qy 167 AlaGlySerProValAsnIleThrLysIleIleTyrglyIleIleIleSerIleMetile 186
Db 522 TCTGTGAGCCGTGTAATATTACGAAGAGATCTTTTGTTCACAAAGTCCATGACATGT 581
Qy 187 ArgThrSerValGlyAsnCys---LysGlnLysGluArgLeuSerValAlaAspAla 205
Db 582 AGATCAGCGTTTGGGCAAGTATTCAAGGAGCAAGCAAAATTTATACAACATAATTAAGAA 641
Qy 206 ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu 225
Db 642 GTTATACTCTTAGCAGGAGGGTTTGATGTGGCTGACATATTCCCTTCATACAAGTCTCT 701
Qy 226 HistTrpIleGlyAlaGlySerLysProArgLeuHisGlnGluIleAspAspIle 245
Db 702 CATGTGCTCAGTGAATGAGGTAAGATTATGAATGCACACCATTAAGTAGATCTATT 761
Qy 246 LeuGluGluIleLeuAsnGluHis-----LysAlaAsnLysPro 258
Db 762 GTTGAGAATGTCATCAACGAGCACAAAGAAATCTTGCAATTTGGGAAACTAATGGAGCG 821
Qy 259 PheGluAlaAspAsnMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValPro 278
Db 822 TTAGAGGTGAAGATTAAATGATCTCTTCTTAAACCTTATGAATGATGGAGGCGCTCAA 881
Qy 279 ValProValThrAsnGlySerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer 298
Db 882 TTTCTATCACCACGACACATCAAGCTATAATCTTTGACATGTTGCTGCTGGAACA 941
Qy 299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
Db 942 GAGACTTCATCGTCAACAATTTGTGGGCTATGTTGGTGAATGTTGTAAGAAATCCAACTGTA 1001
Qy 319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
Db 1002 TTTGCGAAGCTCAAGCAGAGATGAAGATGCAATTTAGAGAAAGAAACCTTTTGATGAA 1061
Qy 339 SerArgPheHisAspLysPheLysLeuValLysGluThrLeuArgLeuHis 358
Db 1062 AATGATGTGGAGAGCTAAACTATCTAAAGTTAGTCATTAAAGAAACTTAAGACTTCAT 1121
Qy 359 ProProVal---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTy 377
Db 1122 CCACCGGTTCCACTTTTGCTCCCAAGAGAATGTAGGGAAGACAAATATAAACGGCTAC 1181
Qy 378 GluIleHisProAsnThrArgIleValValAsnAlaTTPAlaIleGlyArgAspProAsn 397
Db 1182 ACTATTCTGTAAAGACCAAGATCATGTTAATGTTTGGGCATTTGGGAAGAGATCCAAA 1241
Qy 398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTy 417
Db 1242 TATTGGGATGATGCAGAACTTTTAAGCCAGAGAGATTTGACGAGTGCCTTAAGGATT 1301
Qy 418 LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyLe 437
Db 1302 GTTGTAAATATTTTGAATATCTTCATTTGGTGGTGAAGGAGGATTGTCCAGGGATT 1361
Qy 438 ThrSerAlaIleThrAsnLeuGluTyValIleAsnLeuLeuTyHisPheAsnTrp 457
Db 1362 TCGTTTGGTTTGTAGCTAATGCTTATTGTCATTTGGCTCAATTACTTTATCATTGATTGG 1421
Qy 458 GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyAla 477
Db 1422 GAACTCCCACTGGAATCAACCAACGAGCTTTGGACTTGACTGAGTTGGTTGGAGTAAC 1481
```

```
Qy 478 LeuArgLysLysIleAspLeuLysLeuIleProIleProTyrgln 492
Db 1482 GCCGCTAGAAAAAGTGACCTTTACTTGGTTGCGACTCCTTATCAA 1526

RESULT 4
US-10-943-507-153
; Sequence 153, Application US/10943507
; Publication No. US2005013244A1
; GENERAL INFORMATION:
; APPLICANT: U.S. Smokeless Tobacco Company
; TITLE OF INVENTION: 83492-7270 CIP - Cloning of Cytochrome P450 Genes From Nicotiana
; FILE REFERENCE: 83492-7270
; CURRENT APPLICATION NUMBER: US/10/943,507
; PRIOR FILING DATE: 2004-09-17
; PRIOR APPLICATION NUMBER: 10/686,947
; PRIOR FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: 10/387,346
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: 10/340,861
; PRIOR FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 60/363,684
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/347,444
; PRIOR FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 60/337,684
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 153
; LENGTH: 1538
; TYPE: DNA
; ORGANISM: NICOTIANATABACUM
US-10-943-507-153

Alignment Scores:
Pred. No.: 2,77e-121 Length: 1538
Score: 1164.00 Matches: 226
Percent Similarity: 63.4% Conservative: 88
Best Local Similarity: 45.7% Mismatches: 165
Query Match: 45.2% Indels: 16
Dbs: 10 Gaps: 5

US-10-759-813-2 (1-500) x US-10-943-507-153 (1-1538)
```

```
Qy 12 LeuLeSerPheLeuLeuValLeuLeuValValMetArgLeuTrpLysLysGln 31
Db 48 TTGGTTTCCATTTTCCTATTTCTCTTTCTTTTGTAAAGAAATGGAAGAACTCG 107
Qy 32 AsnPro-----ProProGlyProTrpLysPheProIleIleGlyAsnLeu 46
Db 108 AATAGCCAAAGAAAAAATTGCCACCGTCCATGGAAACTTACCANTACTTAGAAGATG 167
Qy 47 ProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle 66
Db 168 CTTCATATGGTT-----GGTGGACTACACACCATGCTCTTAGAGATTTAGCCAAAAA 221
Qy 67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValIleSerSerAla 86
Db 222 TATGACCGCTTATGACCTTTCAATTAGGTGAAGTTTCTGCAAGTTGTGGTTACTTCTCT 281
Qy 87 GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
Db 282 GATATGGCAAGAAAGTAGTATAAAACTCATGACATCGCTTTTCGCTCTTAGGCCTAGCC 341
Qy 107 LeuAspAlaGlnIleValPheTyAsnArgLysAspValLeuPheAlaSerTyrglyAsp 126
Db 342 TTGGCCCGGAGATTGCTGTACATAGGTCTGATCTTCGCTTTGCCCTTATGCGCAT 401
Qy 127 HisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln 146
Db 402 TATTGGACAAATTCGTAATAATATGTCTTCTTGGAAAGTCTCAGTGCACCAAGAAATG 461
Qy 147 SerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLys 166
```

462	ACATAGTAGCTTATTAGCGCGATGAAGGTTCTTCGTCTCTTAATTTATCCGGTCATCT	521
167	AlaGlySerProValAsnIleThrLysIleIleIleIleIleSerIleMetIle	186
522	TCTGGTGAGCCTGTTAATATTACGGAAGGATCTTTTGTTCACAGCTCCATGCATGT	581
187	ArgThrSerValGlyAsnCys---LysGlnLysGluArgLeuLeuSerValAlaAspAla	205
582	AGATCAGCGTTTGGCAAGTATTCAAGGAGCAAGACAAATTTATACAACTAAATTAAGAA	641
206	ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLeuLeu	225
642	GTTATACTTTAGCAGGAGGTTTGATGTGGCTGACATATTCCTTCATACAGTCTCT	701
226	HisTyrlleIleGlyAlaGluSerLysProArgGluLeuHisGlnGluIleAspIle	245
702	CATGTCTCAGTGAATCAAGGTAAGATTATGAATGCACACCATAAGGTAGTCTATT	761
246	LeuGluGluLeuLeuAsnGluHis-----LysAlaAsnLysPro	258
762	GTTGAGAATGTTCATCAACGAGCAACAGAAAATCTTGCAATTGGGAAAACTAATGGAGCG	821
259	PheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValPro	278
822	TTAGGAGGTGAAGATTTAAATTGATGTTCTTCAAACTTATGAATGATGGAGCGCTTCAA	881
279	ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer	298
882	TTTCCTATCAACACGACATCAAGCTATATACTTTTGACATGTTGTCTGGTGAACA	941
299	GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu	318
942	GAGACTTCATCGTCAACAAATTGTGTGGCTATGTGGGAATGGTGAAAAATCCAACGTGA	1001
319	LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu	338
1002	TTTTCGGAAGCTCAAGCAGAAGTAAGAGATGCATTTTAGAGAAAAAGAAACTTTTGATGA	1061
339	SerArgPheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHis	358
1062	ANTGATGTGGAGGACTAACTACTAAAGTTAGTCAITTAAGAACTCTAAGACTTCAT	1121
359	ProProVal---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyr	377
1122	CCACCGTTCCACTTTTGCTCCCAAGAGAATGTAGGGAAGAGACAAATATAACGGCTAC	1181
378	GluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsn	397
1182	ACTATTCCTGTAAAGACCAAGTCATGGTTAAATGTTTGGGCATTGGGAAGAGATCCA	1241
398	ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr	417
1242	TATTGGATGATGCAGAAACTTTTAACCCAGAGAGATTGACGACGTCTCTAAGGATTTT	1301
418	LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIle	437
1302	GTTGGTAAATATTTGAATATCTTCCATTTGGTGGTGAAGGAGGATTTGTCACGGATT	1361
438	ThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrlleHisPheAsnTrp	457
1362	TCGTTTGGTTAGCTAATGCTTATTTGCCATTGGCTCAATTAATCTTATCACTTTGATTGG	1421
458	GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAla	477
1422	GAACTCCCCACTGGAATCAACACCAAGCGACTTGGACTTGAATGGTGGTGGAGTAACT	1481
478	LeuArgLysLysIleAspLeuLysLeuIleProIleProTyrGln	492
1482	CGCGCTAGAAAAAGTGACCTTTACTTGGTGGGACTCCTTATCA	1526

RESULT 5

US-10-934-944-153

```

; Sequence 153, Application US/10934944
; Publication No. US20060037096A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Dongmei
; TITLE OF INVENTION: Cloning of Cytochrome P450 Genes From Nicotiana
; FILE REFERENCE: 07678/141008
; CURRENT APPLICATION NUMBER: US/10/934,944
; CURRENT FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: 10/686,947
; PRIOR FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: 60/503,989
; PRIOR FILING DATE: 2003-09-16
; PRIOR APPLICATION NUMBER: 60/485,368
; PRIOR FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: 10/387,346
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: 10/340,861
; PRIOR FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 10/293,252
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 60/418,933
; PRIOR FILING DATE: 2002-10-16
; PRIOR APPLICATION NUMBER: 60/363,684
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/347,444
; PRIOR FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 60/337,684
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 387
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 153
; LENGTH: 1538
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
US-10-934-944-153

Alignment Scores:
Pred. No.: 2.77e-121 Length: 1538
Score: 1164.00 Matches: 226
Percent Similarity: 63.4% Conservative: 88
Best Local Similarity: 45.7% Mismatches: 165
Query Match: 45.2% Indels: 16
DB: 11 Gaps: 5

US-10-759-813-2 (1-500) x US-10-934-944-153 (1-1538)

Qy 12 LeuLeSerPheLeuLeuValValLeuValValMetArgLeuTrpLysLysGln 31
Db 48 TTGGTTCCCATTTTCCTATCTTCTATCTTTTCTCTTTTAAAGGAAATGGAAGAACTCG 107

Qy 32 AsnPro-----ProGlyProTrpLysPheProIleLeGlyAsnLeu 46
Db 108 ANTAGCCAAAGAAAAATTCACACAGTTCATGGAACCTACCAATACCTAGGAAGTAGT 167

Qy 47 ProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle 66
Db 168 CTTTCATATGGTT-----GGTGGACTACACACCATGCTCTTAGAGATTTAGCCAAAAA 221

Qy 67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSerAla 86
Db 222 TATGGACCGCTTATGACCTTCAATATAGGTGAAGTTTCTGCAGTGTGTGCTTACTTCTCCT 281

Qy 87 GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
Db 282 GATATGGCAAAAGAGTACTAAAAACTCATGACATCGCTTTTCGCGTCTAGGCCCTAGCCCT 341

Qy 107 LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
Db 342 TTGGCCCGGAGATTGCTGTTCACATAGGCTGTGATCTTCGCTTTTGGCCCTATGGCGAT 401

Qy 127 HisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln 146
Db 402 TATTGGAGACAAATGCGTAAATATGCTCTTGGAGTGCTCAGTGCCAAAGATGTTCGG 461

```



```
QY 206 ValAenGluAlaAlaThrSerPheGlyThrAlaAaspAlaPheProThrTrpLysLeuLeu 225
Db 648 GTGATACCTCTTAGCAGGAGGGTTGATGTGGCTGACATATTCCTTCTACTGAAGTTCTT 707
QY 226 HistyrIleIleGlyAlaGluSerLysProArgArgLeuHisGlnGluIleAaspAspIle 245
Db 708 CATGTGCTCAGTGGAAATGAAGGGTAAGATTATGAATGCACACCATATAGGTAGATGCCATT 767
QY 246 LeuGluGluIleLeuAasnGluHis-----LysAlaAasnLysPro 258
Db 768 GTTGAGATGTCATCAATGAGCACAAGAAAAATCTTGCAATTTGGGAAAACTAATGGAGCG 827
QY 259 PheGluAlaAaspAenLeuMetAaspValLeuLeuAenLeuGlnLysAasnGlyAenValPro 278
Db 828 TTAGGAGGTGAAGATTAAATGATGTTCTTCTTAAGACTTTATGAATGATGAGGCGCTTCAA 887
QY 279 ValProValThrAasnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer 298
Db 888 TTTCCTATCACCACGACATCAAGCCATAATTTTGGACATGTTTGCTGCCGGGACA 947
QY 299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAasnProThrGlu 318
Db 948 GAGACTTCATCTCAACAATTGTGTGGCTATGTTAGTAAATCGTGAAAAATCCAGCCGTA 1007
QY 319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAaspGlu 338
Db 1008 TTCCGGAAGCTCAAGCAGCAAGTAAGAGAACATTTAGAGGAAAAAGAACTTTTCGATGAA 1067
QY 339 SerArgPheHisAaspLeuLysPheLysLeuValValLysGluThrLeuArgLeuHis 358
Db 1068 AATGATGTGGAGGAGCTAACTACCTAAAGTTAGTAAATAAAGAACTCTAAGACTTCAT 1127
QY 359 ProProVal---ValLeuIleProArgGluCysArgGluThrThrArgIleAaspGlyTyr 377
Db 1128 CCACCGGTCCACTTTGCTCCAGAGAAATGTAGGGAAGACAAATAATAACCGCTAC 1187
QY 378 GluIleHisProAenThrArgIleValValAenAlaTrpAlaIleGlyArgAaspProAen 397
Db 1188 ACTATTCTCTGTAAGACCAAAAGTCATGTTAATGTTTGGGCTTTGGGAAGAGATCCAAA 1247
QY 398 ThrTrpSerGluProGlyLysPheAenProGluArgPheLysAaspCysAlaIleAaspTyr 417
Db 1248 TATTGGAATGACGCAAACTTTTATGCCAGAGATTTTGACATGCTCTAAGGATTTT 1307
QY 418 LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIle 437
Db 1308 GTTGGTAATAATTTGAATATCTTCCATTGTGTGGCGGAAGGAGATTGTCTCGGATT 1367
QY 438 ThrSerAlaIleThrAenLeuGluTyrValIleIleAenLeuLeuTyrHisPheAenTrp 457
Db 1368 TCGTTTGGCTTAGCTAATGCTTTATTGGCAATGGCTCAATTACTATATCACTTCGATTGG 1427
QY 458 GluLeuAlaAaspGlyIleThrProGlnThrLeuAaspMetThrGluAlaIleGlyAla 477
Db 1428 AAATCCCTGCTGGAAATCGAACCAAGCGACTTGGACTTGACTGAGTTGGTTGGAGTAAC 1487
QY 478 LeuArgLysLysIleAaspLeuLysLeuIleProIleProTyrGln 492
Db 1488 GCCGCTAGAAAAAGTGACCTTTACTTGGTTGGCGACTCTTATCAA 1532
```

RESULT 8

```
US-10-686-947-231
; Sequence 231, Application US/10686947
; Publication No. US20040162420A1
; GENERAL INFORMATION:
; APPLICANT: Profigen Inc.
; TITLE OF INVENTION: Cloning of Cytochrome P450 Genes from Tobacco
; FILE REFERENCE: 79601
; CURRENT APPLICATION NUMBER: US/10/686,947
; CURRENT FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: US 10/387346
; PRIOR FILING DATE: 2003-03-12
```

```
; NUMBER OF SEQ ID NOS: 298
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 231
; LENGTH: 1673
; TYPE: DNA
; ORGANISM: NICOTIANA TABACUM
US-10-686-947-231
```

Alignment Scores:

Pred. No.:	1.62e-119	Length:	1673
Score:	1149.00	Matches:	222
Percent Similarity:	63.2%	Conservative:	91
Best Local Similarity:	44.8%	Mismatches:	166
Query Match:	44.8%	Indels:	16
DB:	8	Gaps:	5

US-10-759-813-2 (1-500) x US-10-686-947-231 (1-1673)

```
QY 12 LeuIleSerPheLeuLeuValLeuIleLeuValValMetArgLeuTrpLysLysGln 31
Db 54 TTGGTTTCCATTTTCCTTATTTCTTCTTCTTTTGGTATGATGAGAACTCC 113
QY 32 AasnPro-----ProProGlyProTrpLysPheProIleIleGlyAenLeu 46
Db 114 AATAGCCAAAGCAAAAAAGTTGCCACCAAGTCCATGGAAACTACCAATACTAGGAAGTATG 173
QY 47 ProHisLeuLeuLeuThrSerAaspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle 66
Db 174 CTTCATATGGTT-----GGTGGACTACCAACCATGTCCTTAGAGATTTAGCCAAAAA 227
QY 67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSerAla 86
Db 228 TATGACCACTTATGACCTTCAATTAGTGAAGTTTCTGCGTTGTGGTTACTTCTCT 287
QY 87 GluAlaAlaLysGluValMetLysThrGlnAlaAaspAlaPheAlaGlnArgProIleVal 106
Db 288 GATACGGCAAAAGAGATTATTAANAACATCATGACATCGCTTTTGGCTTAGGCCTAGCCTT 347
QY 107 LeuAaspAlaGlnIleValPheTyrAasnArgLysAaspValLeuPheAlaSerTyrGlyAasp 126
Db 348 TTGGCCCGGAGATTGCTGTACAAATAGGTCTGTATCTAGCCTTTTGGCCCTATGCGCAC 407
QY 127 HistTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln 146
Db 408 TATTGGAGCAAAATCGTAAATAATGTCTTGTGGAAGTGTCTCAGTCCCAAGAAATTTCCG 467
QY 147 SerSerArgLeuIleAargGluGluMetGluAaspAlaIleThrPheLeuArgSerLys 166
Db 468 ACATTTAGCTCTATTAGCGGAATGAAGTTCTTCTGCTCTCAATTAATTTATCGGTCACT 527
QY 167 AlaGlySerProValAenIleThrLysIleIleTyrGlyIleIleIleSerIleMetIle 186
Db 528 TCTGTGTAACCTATTAATGTTACGGAAGAGGATCTTTTGTTCACAAGTCCATGACATGT 587
QY 187 ArgThrSerValGlyAasnCys---LysGlnLysGluArgLeuLeuSerValAlaAaspAla 205
Db 588 AGATCAGCGTTTGGCAAGTGTTCNAAGAGCAAGCAAAATTTATACAACATAATTAAGAA 647
QY 206 ValAenGluAlaAlaThrSerPheGlyThrAlaAaspAlaPheProThrTrpLysLeuLeu 225
Db 648 GTGATACTCTTAGCAGGAGGGTTTGATGTGGCTGACATATTCCTTCACTGAAGTTCTT 707
QY 226 HistyrIleIleGlyAlaGluSerLysProArgArgLeuHisGlnGluIleAaspAspIle 245
Db 708 CATGTGCTCAGTGGAAATGAAGGTAAAGATTATGAATGCACACCATTAAGTAGATGCCATT 767
QY 246 LeuGluGluIleLeuAasnGluHis-----LysAlaAasnLysPro 258
Db 768 GTTGAGATGTCATCAATGAGCACAAGAAAAATCTTGCAATTTGGGAAAACTAATGGAGCG 827
QY 259 PheGluAlaAaspAenLeuMetAaspValLeuLeuAenLeuGlnLysAasnGlyAenValPro 278
Db 828 TTAGGAGGTGAAGATTAAATGATGTTCTTCTTAAGACTTTATGAATGATGAGGCGCTTCAA 887
```



```
QY 279 ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer 298
Db 888 TTTCTCTATCACCACACACACATCAAGCATATTTTTCACATGTTTGGTCGGCGGACA 947
QY 299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
Db 948 GAGACTTTCATCGTCAACAATTGTGTGGCTATGCTAGAAATGGTGAATAATCCAGCGGTA 1007
QY 319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
Db 1008 TTCCGGAAGCTCAAGCAGAAAGTAAGAGACATTTAGAGAAAGAACTTTCCGATGAA 1067
QY 339 SerArgPheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHis 358
Db 1068 AATGATGTGGAGGAGCTAAACTACCTAAAGCTTAGTAATAAAGAACTCTAAGACTTCAT 1127
QY 359 ProProVal---ValLeuIleProArgGluCysArgGluThrThrArgLysAspGlyTyr 377
Db 1128 CCACCGGTTCCACTTTTGTCTCCCAAGAGATGTAGGGAAGACAAATATAAACCAGCTAC 1187
QY 378 GluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsn 397
Db 1188 ACTATTCTCTGTAAGACCAAGCATGGTTAATGTTGGGCTTTGGGAAGAGATCCAAAA 1247
QY 398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr 417
Db 1248 TATTGGAATGACGCAAACTTTATGCCAGAGAGATTGAGCAGTGCTCTAAGGATTTT 1307
QY 418 LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgLysCysProGlyIle 437
Db 1308 GTTGGTAATAATTTTGAATATCTTCCATTTGGTGGCGAAGAGGATTTGCTCGGATT 1367
QY 438 ThrSerAlaIleThrAsnLeuGluTyrValIleAsnLeuLysThrHisPheAsnTrp 457
Db 1368 TCGTTGGCTTAGCTAATGCTTATTGTCATGCTGCTCAATATATATCATCTTCGATTGG 1427
QY 458 GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAla 477
Db 1428 AACTCCCTCGTGGATCGAATCGAACCAGGACTTGGACTTGCATGCTGGTGGAGTAAC 1487
QY 478 LeuArgLysLysIleAspLeuLysLeuIleProIleProTyrGln 492
Db 1488 GCGCTAGAAAAAGTGACCTTTACTTTGGTTGGGACTCTCTATCA 1532

RESULT 9
US-10-943-507-208
; Sequence 208, Application US/10943507
; Publication No. US2005013244A1
; GENERAL INFORMATION:
; APPLICANT: U.S. Smokeless Tobacco Company
; TITLE OF INVENTION: 83492-7270 CIP - Cloning of Cytochrome P450 Genes From Nicotiana
; FILE REFERENCE: 83492-7270
; CURRENT APPLICATION NUMBER: US/10/943,507
; CURRENT FILING DATE: 2004-09-17
; PRIOR APPLICATION NUMBER: 10/686,947
; PRIOR FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: 10/387,346
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: 10/340,861
; PRIOR FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 60/363,684
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/347,444
; PRIOR FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 60/337,684
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 208
; LENGTH: 1673
; TYPE: DNA
; ORGANISM: NICOTIANATABACUM
```

US-10-943-507-208

```
Alignment Scores:
Pred. No.: 1,62e-119 Length: 1673
Score: 1149.00 Matches: 222
Percent Similarity: 63.2% Conservative: 91
Best Local Similarity: 44.8% Mismatches: 166
Query Match: 44.8% Indels: 16
DB: 10 Gaps: 5
```

US-10-759-813-2 (1-500) x US-10-943-507-208 (1-1673)

```
QY 12 LeuIleSerPheLeuLeuValLeuIleLeuValValMetArgLeuTrpLysLysGln 31
Db 54 TTGGTTTCCATTTCTCTATTTCTATCTTTCTCTTTTGTAGGATATAGGAAGAACTCC 113
QY 32 AsnPro-----ProProGlyProTrpLysPheProIleLeuGlyAsnLeu 46
Db 114 AATAGCCAAAGCAAAAGTTGCCACCGAGTCCATGGAAACTACCAATACTAGGAAGTATG 173
QY 47 ProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle 66
Db 174 CTTCTATGCTGTT-----GGTGGACTACCAACCATGTCCTTAGAGATTTAGCCAAAAA 227
QY 67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSerAla 86
Db 228 TATGGACCACTTATGACCTTCAATTAGGTGAAGTTTCTGCGGTTGTGGTTACTTCTCT 287
QY 87 GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
Db 288 GATACGCCAAAGAGAGTATTAATAAAGTCTATGACATCGCTTTTGGCGTCTAGGCTT 347
QY 107 LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
Db 348 TTGGCCCCGGAGATTGTCTGTACAAATAGGTCTGATCTAGCTCTTGGCCCTATGGCGAC 407
QY 127 HistTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln 146
Db 408 TATTGGAGACAAATCGGTAAATATGCTGTCTTGGGAAGTCTCAGTGCACCAAGATGTT 467
QY 147 SerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLys 166
Db 468 ACATTTAGCTCTATTAGGCGGAATGAAGTCTTCGTCTCATTAATTTATTCGGTCTAT 527
QY 167 AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleSerIleMetIle 186
Db 528 TCTGTGAACCTATTAAATGTTACGGAAGGATCTTTTGTTCACAAGCTCCATGACATGT 587
QY 187 ArgThrSerValGlyAsnCys---LysGlnLysGluArgLeuLeuSerValAlaAspAla 205
Db 588 AGATCAGCGTTTGGGCAAGTGTTCAAAGAGCAAGACAAATTTATACAACCTAATTAAGAA 647
QY 206 ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu 225
Db 648 GTGATACTCTTAGCAGGAGGGTTTGATGTGGCTGACATATTCCTTCTACTGAAGTTTCT 707
QY 226 HistTyrIleIleGlyAlaGluSerLysProArgLysGlnLeuIleAspAspIle 245
Db 708 CATGTGTCAGTGAATGAAGGTATGATATGACATGACACCATTAAGGTAGTAGCATTT 767
QY 246 LeuGluGluIleLeuAsnGluHis-----LysAlaLeuLysPro 258
Db 768 GTTGAGATGTCATCAATGAGCACAAGAAAAATCTTGCATTTGGGAAAAAATAATGAGCG 827
QY 259 PheGluAlaAspAsnLeuMetAspValLeuLeuLeuAsnLeuGlnLysAsnGlyAsnValPro 278
Db 828 TTAGGAGGTGAAGATTAATTAATGATGTTCTTCTTAAGACTTATGAATGATGAGGCGCTCAA 887
QY 279 ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrIleGlySer 298
Db 888 TTTCTATCACCAACGACAAATCAAGCCATAAATTTTGGACATGTTTGTCTCCCGGACACA 947
QY 299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
```



```
Db 948 GAGACTTCATCGTCAACAAATTGTGGGCTATGGTAGAAATGGTGAATAATCCAGCCGCTA 1007
Qy 319 LeuArgLysAlaGlnGluGlnValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
Db 1008 TTCGGGAAGCTCAGCAGAGTAGAGAGCAATTTAGAGGAAAGAAACTTTTCGATGAA 1067
Qy 339 SerArgPheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHis 358
Db 1068 AATGATGTGGAGGAGTAACTACCTAACCTACCTAAGATTAGTAAATAAAGAACTCTAAGACTTTCAT 1127
Qy 359 ProProVal---ValLeuLeuProArgGluCysArgGluThrThrArgIleAspGlyTyr 377
Db 1128 CCACCGGTTCCACTTTTGTCTCCAGAGAAATGTAGGGAAGACAAATATAAACGGCTAC 1187
Qy 378 GluIleHisProAsnThrArgIleValValAsnAlaTtpAlaIleGlyArgAspProAsn 397
Db 1198 ACTATTCCTGTAAAGCAACCAAGTCATGTTAATGTTTGGCTTTGGGAAGAGATCCAAA 1247
Qy 398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr 417
Db 1248 TATTGGAATGACCGCAAACTTTTATGTCAGAGAGATTGAGCAGTGTCTTAAGGATTTT 1307
Qy 418 LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIle 437
Db 1308 GTTGTAATAATTTTGAATATCTTCATTTGGTGGGAAGAGATTGCTCGGGATT 1367
Qy 438 ThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrp 457
Db 1368 TCGTTTGGCTTAGCTAATGCTTATTTGCCATTGGCTCAATTACTATATACCTTCGATTGG 1427
Qy 458 GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyAla 477
Db 1428 AAACTCCCTGCTGGAATCAACCAAGCGACTTGGACTTGAGTGTGTTGGAGTAAC 1487
Qy 478 LeuArgLysLysIleAspLeuLysLeuIleProIleProTyrGln 492
Db 1488 GCCGCTAGAAAAAGTGACCTTTACTTGGTTGGCACTCTTATCAA 1532
```

RESULT 10

```
US-10-934-944-211
; Sequence 211, Application US/10934944
; Publication No. US20060037096A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Dongmei
; TITLE OF INVENTION: Cloning of Cytochrome P450 Genes From Nicotiana
; FILE REFERENCE: 07678/141008
; CURRENT APPLICATION NUMBER: US/10/934,944
; CURRENT FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: 10/686,947
; PRIOR FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: 60/503,989
; PRIOR FILING DATE: 2003-09-18
; PRIOR APPLICATION NUMBER: 60/485,368
; PRIOR FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: 10/387,346
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: 10/340,861
; PRIOR FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 10/293,252
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 60/418,933
; PRIOR FILING DATE: 2002-10-16
; PRIOR APPLICATION NUMBER: 60/363,684
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/347,444
; PRIOR FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 60/337,684
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 387
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 211
; LENGTH: 1673
```

```
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
US-10-934-944-211
Alignment Scores:
Pred. No.: 1,62e-119 Length: 1673
Score: 1149.00 Matches: 222
Percent Similarity: 63.2% Conservative: 91
Best Local Similarity: 44.8% Mismatches: 166
Query Match: 44.6% Indels: 16
DB: Gaps: 5
US-10-759-813-2 (1-500) x US-10-934-944-211 (1-1673)
Qy 12 LeuIleSerPheLeuLeuValValLeuValValMetArgLeuTrpLysLysGln 31
Db 54 TTGGTTCCATTTCTTCTATCTTCTCTCTTTTGTAAAGGATATGAAGAACAATCC 113
Qy 32 AsnPro-----ProGlyProTrpLysPheProIleIleGlyAsnLeu 46
Db 114 AATAGCCAAACAAAAAGTTGCCACCAAGTCCATGGAAACTTACCNAATCTAGGAAGTAG 173
Qy 47 ProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle 66
Db 174 CTTCATATGGTT-----GGTGGACTACCAACCATGCTCTTAGAGATTTAGCCAAAAA 227
Qy 67 TyrGlyProValMetSerLeuGlnIleGlyValSerAlaValValIleSerSerAla 86
Db 228 TATGACCACTTATGACCTTCAATTAGGTGAAGTTCTCGCGTTGTGGTTACTTCTCT 287
Qy 87 GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
Db 288 GATACGGCAAAAGAGATTTAAAAACTCATGACATCGCTTTTGGCTAGCGCTAGCCTT 347
Qy 107 LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
Db 348 TTGCCCCGGAGATTGCTGTACAAATAGGTCTGATCTAGCCTTTTGGCCCTATGCGCAC 407
Qy 127 HisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln 146
Db 408 TATTGGAGACAAATGCGTAAATATGTCTTGGAAAGTGCCTAGTGCACAAAGATTTTCGG 467
Qy 147 SerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLys 166
Db 468 ACATTTAGCTCTATTAGCGGAATGAAGTTCTTGTCTCATTAATTTATTCGGTCATCT 527
Qy 167 AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleIleSerIleMetIle 186
Db 528 TCTGTGAACCTATTATATGTTACGGAAGGATCTTTTGTTCACAAAGCTCCATGACATGT 587
Qy 187 ArgThrSerValGlyAsnCys---LysGlnLysGluArgLeuLeuSerValAlaAspAla 205
Db 588 AGATCAGCGTTTGGCAAGTGTTCAAAGAGCAAGACAAATTTATACAACTAATTAAGAA 647
Qy 206 ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu 225
Db 648 GTGATACTCTTAGCAGGAGGGTTTGATGTGCTGCATATATCCCTTCTACTGAAGTTCTT 707
Qy 226 HistyTrpIleIleGlyAlaGluSerLysProArgArgLeuHisGlnGluIleAspIle 245
Db 708 CATGTGCTCAGTGGGAATGAAGGTAAAGATTATGAATGCACACCAATAAGTAGATGCCATT 767
Qy 246 LeuGluGluIleLeuAsnGluHis-----LysAlaAsnLysPro 258
Db 768 GTTGAGAAATGTCATCAATGAGCACAAAGAAATCTTGGCAATTTGGGAAACTAATGGAGCG 827
Qy 259 PheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValPro 278
Db 828 TTAGGAGGTGAAGATTTAATTGATGTTCTTCTAAGACTTATGAATGATGAGGAGCCTTCAA 887
Qy 279 ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrIleGlySer 298
Db 888 TTTCCTATCACCAACGACAAACATCAAGCCATAATTTTGTGACATGTTTGTCTCCGGGACA 947
```



```
Db      828 TTAGGAGGTGAAGATTAATTTGATGTTCTTCTAAGACTTATGAATGATGAGCGCTTCAA 887
Qy      279 ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer 298
Db      888 TTTCCTATACCAACAGACATCAAGCCATTAATTTTGCATGTTTGTGCTCGGAGCA 947
Qy      299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
Db      948 GAGACTTCATCGTCAACAATTTGTGGCTATGTTAGAAATGTTGAAATATCCAGCCGTA 1007
Qy      319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
Db      1008 TTCGCGAAAGCTCAAGCAGAGTAAGAGAGACATTTAGAGGAAAGAACTTTTCGATGAA 1067
Qy      339 SerArgPheHisAspLeuLysPhePheLysLeuValLysGluThrLeuArgLysHis 358
Db      1068 AATGATGTGGAGAGCTAACTACCTAAAGTTAGTATATAAAGAACTCTAAGACTTCAT 1127
Qy      359 ProProVal---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyr 377
Db      1128 CCACCGGTTCCACTTTTGTCTCCCAAGAGATGTAGGAGAGACAAATATAACGCTAC 1187
Qy      378 GluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsn 397
Db      1188 ACTATTTCCTGTAAGACCAAGATCATGTTTAATGTTTGGCTTTTGGGAAGAGATCCAAA 1247
Qy      398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr 417
Db      1248 TATTGGAATGACCGCAAACTTTTATGCCAGAGAGATTTGAGCAGTGCTCTAAGGATTTT 1307
Qy      418 LysGlyThrThrPheGluLeuValProPheGlyValAlaGlyLysArgIleCysProGlyIle 437
Db      1308 GTTGGTAATAATTTTGAATATCTTCCATTTGTGGCGGAGAGAGATTGCTCTGGGATT 1367
Qy      438 ThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrp 457
Db      1368 TCGTTGGCTTAGCTAATGCTTATTTGGCATTTGGCTCAATTTACTATATCACTTCGATTGG 1427
Qy      458 GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyAla 477
Db      1428 AAATCCCTCGTGGATCGAACCAAGCGACTTGGACTTGGACTTGGATTGGTGGAGTAAC 1487
Qy      478 LeuArgLysLysIleAspLeuLysLeuIleProIleProTyrGln 492
Db      1488 GCGGCTAGAAAAGTGCACCTTTACTTGTGGTGGAGCTCTTATCA 1532
```

RESULT 13

```
US-11-116-881A-240
; Sequence 240, Application US/11116881A
; Publication No. US20060041949A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Dongmei
; APPLICANT: Nielsen, Mark T.
; TITLE OF INVENTION: Nicotiana Nucleic Acid Molecules and Uses Thereof
; FILE REFERENCE: 07678/141014
; CURRENT APPLICATION NUMBER: US/11/116,881A
; CURRENT FILING DATE: 2005-04-27
; PRIOR APPLICATION NUMBER: 60/665,451
; PRIOR FILING DATE: 2005-03-24
; PRIOR APPLICATION NUMBER: 60/665,097
; PRIOR FILING DATE: 2005-03-24
; PRIOR APPLICATION NUMBER: 60/646,764
; PRIOR FILING DATE: 2005-01-25
; PRIOR APPLICATION NUMBER: 60/607,357
; PRIOR FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: 60/566,235
; PRIOR FILING DATE: 2004-04-29
; PRIOR APPLICATION NUMBER: 10/934,944
; PRIOR FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: 10/943,507
; PRIOR FILING DATE: 2004-09-17
; PRIOR APPLICATION NUMBER: 60/503,989
```

```
; PRIOR FILING DATE: 2003-09-18
; PRIOR APPLICATION NUMBER: 60/485,368
; PRIOR FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: 60/418,933
; PRIOR FILING DATE: 2002-10-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2300
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 240
; LENGTH: 1673
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
US-11-116-881A-240
```

Alignment Scores:

Pred. No.:	1,62e-119	Length:	1673
Score:	1149.00	Matches:	222
Percent Similarity:	63.2%	Conservative:	91
Best Local Similarity:	44.8%	Mismatches:	166
Query Match:	44.6%	Indels:	16
DB:	16	Gaps:	5

US-10-759-813-2 (1-500) x US-11-116-881A-240 (1-1673)

```
Qy      12 LeuIleSerPheLeuLeuValLeuIleLeuValValMetArgLeuTrpLysLysGln 31
Db      54 TTGTTTCCATTTTCCTATTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 113
Qy      32 AsnPro-----ProGlyProTrpLysPheProIleIleGlyAsnLeu 46
Db      114 AATAGCAAAAGCAAAAGTTGCCACGAGTCCATCGAAACTACCAATACTAGGAAGTATG 173
Qy      47 ProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIle 66
Db      174 CTTTCATATGGT-----GGTGGACTACCCACCATGTCTCTTAGAGATTTAGCCAAA 227
Qy      67 TyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSerAla 86
Db      228 TATGACCACTTATGACCTTCAATTAGGTGAAGTTCTTCGGTCTGCTGTTACTTCTCTCT 287
Qy      87 GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVal 106
Db      288 GATACGGCAAAAAGAGATATTAAAACTCATGACATCGCTTTTTCGCTCTAGGCTAGCCTT 347
Qy      107 LeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
Db      348 TTGGCCCGGAGATTGCTGTGTACAAATAGTCTGTATCTAGCTTTTTCGCCCTTAGCGGAC 407
Qy      127 HisTrpArgGlnMethLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGln 146
Db      408 TATTGGAGACAAATCGTAAATATATGCTTCTTGGAGTGTCTCAGTCCCAAGATGTTCCG 467
Qy      147 SerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLys 166
Db      468 ACATATTAGCTCTATTAGGCGGAATGAAGTTCTTCGTCTCATTAATTTTATCCGTCATCT 527
Qy      167 AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleIleSerIleMetIle 186
Db      528 TCTGTGAACCTATTAAATGTTACGAAAGAGATCTTTTTCACAAAGCTCCATGACATGT 587
Qy      187 ArgThrSerValGlyAsnCys---LysGlnLysGluArgLeuLeuSerValAlaAspAla 205
Db      588 AGATCAGCGTTTGGGCAAGTGTCAAGAGCAAGCAAAATTTATACAATTAATTAAGAA 647
Qy      206 ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu 225
Db      648 GTGATACCTCTTAGCAGAGGGTGTGATGTGGCTGACATATTCCCTTCACTGAAGTTCTT 707
Qy      226 HisTyrIleIleGlyAlaGluSerLysProArgLeuHisGlnIleAspIle 245
Db      708 CATGTGCTCAGTGAATGAAGGTGAATATGAATATGACACCAATGAAGGTAGATGCCATT 767
Qy      246 LeuGluGluIleLeuAsnGluHis-----LysAlaAsnLysPro 258
```



```

1008 TTCCGCGAAAGCTCAAGCAGAAGTAAGAGAGCATTTTAGAGGAAAGAACTTTTCGATGAA 1067
339 SerArgPheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHis 358
1068 AATGATGTGAGGAGCTAACTACCTAAAGTTAGTAATAAAGAAACTCTAAGACATTCAT 1127
359 ProProVal---ValLeuLeuProArgGluCyAsArgGluThrThrArgLysAspGlyTyr 377
1128 CCACCGGTTCCACTTTTGCTCCCAAGAGAATGTAGGGAAGAGACAATATAACGGCTAC 1187
378 GluHisProAsnThrArgLysLeuValValAsnAlaTirPalalleGlyArgAspProAsn 397
1188 ACTATTCCCTGTAAGACCAAGAGTCATGGTTAATGTTTGGGCTTTGGGAAGAGATCCAAAA 1247
398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlalleAspTyr 417
1248 TATTGGAATGACCGAAGAACTTTATGCCAGAGAGATTTGAGCAGTGCTCTAAGGATTTT 1307
418 LysGlyThrThrPheGluLeuValProPheGlyValaGlyLysArgLysCysProGlyLys 437
1308 GTTGGTAATAATTTTGAATATCTTCCATTTGGTGGCGGAAGAGGATTTGCTCTGGGATT 1367
438 ThrSerAlalleThrAsnLeuGluTyrValLilleAsnLeuLeuTyrHisPheAsnTrp 457
1368 TCGTTTGGCTTAGCTAATGCTAATTTTGGCATTTGGCTCAATTTACTATATACATTCGATTGG 1427
458 GluLeuAlaAspGlyLysLeuThrProGlnThrLeuAspMetThrGluAlalleGlyAla 477
1428 AAACCTCCCTGCTGGAATCGAACCAGCGACTTGGACTTGGACTTGGACTTGGACTTGGACT 1487
478 LeuArgLysLysLysLeuAspLeuLysLeuLysLeuLysLeuLysLeuLysLeuLysLeu 492
1488 GCGGCTAGAAAAGTGACCTTTACTTTGGTTGGGACTTCTTAACA 1532

RESULT 15
US-10-943-507-206
; Sequence 206, Application US/10943507
; Publication No. US20050132444A1
; GENERAL INFORMATION:
; APPLICANT: U.S. Smokeless Tobacco Company
; TITLE OF INVENTION: 83492-7270 Clip - Cloning of Cytochrome P450 Genes From Nicotiana
; FILE REFERENCE: 83492-7270
; CURRENT APPLICATION NUMBER: US/10/943,507
; CURRENT FILING DATE: 2004-09-17
; PRIOR APPLICATION NUMBER: 10/686,947
; PRIOR FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: 10/387,346
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: 10/340,861
; PRIOR FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 60/363,684
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/347,444
; PRIOR FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 60/337,684
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 206
; LENGTH: 1673
; TYPE: DNA
; ORGANISM: NICOTIANA TABACUM
US-10-943-507-206

Alignment Scores:
Pred. No.: 21e-119 Length: 1673
Score: 1148.00 Matches: 222
Percent Similarity: 63.2% Conservative: 91
Best Local Similarity: 44.8% Mismatches: 166
Query Match: 44.6% Indels: 16
DB: 10 Gaps: 5
```

```

US-10-759-813-2 (1-500) x US-10-943-507-206 (1-1673)
QY 12 LeuLeuSerPheLeuLeuValLeuValValMetArgLeuTyrLysLysGln 31
Db 54 CTGGTTTCCATTTTCTTATCTTTCCTCTTCTTAAAGGATATGGAAGAACTCC 113
QY 32 AsnPro-----ProProGlyProTyrLysPheProLysLysLysLys 46
Db 114 AATAGCCAAAGCAAAAGTTGCCACCAAGTCCATGGAACTACCAATACCTAGGAAGTATG 173
QY 47 ProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnLe 66
Db 174 CTTTCATATGGTT-----GGTGGACTACCAACCATGTCTCTTAGAGATTTAGCCAAAAA 227
QY 67 TyrGlyProValMetSerLeuGlnLeuGlnValSerAlaValValLysSerSerAla 86
Db 228 TATGGACCACTTATGACCTTCAATTAGGTGAAGTTTTCGGGTTTGGTTACTTCTTCCT 287
QY 87 GluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProLysVal 106
Db 288 GATACGGCAAAAGAGTATTAATAAACTCATGACATCGCTTTTGGCTAGGCCCTAGCCCT 347
QY 107 LeuAspAlaGlnLeuValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAsp 126
Db 348 TTGGCCCGGAGATTGCTGTACAAATAGTCTGTAGCTAGCTTTTGGCCCTATGGCGAC 407
QY 127 HisTrpArgGlnMetLysLysLysLysLysLysLysLysLysLysLysLysValGln 146
Db 408 TATTTGAGACAATCGTAAATAATGTGTCTTGGAAAGTGTCTCAGTCCCAAGAAATGTCGG 467
QY 147 SerSerArgLeuLeuArgGluGluMetGluAspAlaLysLeuThrPheLeuArgSerLys 166
Db 468 ACATTTAGCTCTATTAGCGGAATGAAGTTCTTCTGCTCATTAATTTATCCGTCATCT 527
QY 167 AlaGlySerProValAsnLysLysLysLysLysLysLysLysLysLysLysLysMetIle 186
Db 528 TCTGTGGAACCTATTAAATTTACGGAAGGATCTTTTGTTCACAAAGCTCCATGACATGT 587
QY 187 ArgThrSerValGlyAsnCys---LysGlnLysGluArgLeuLeuSerValAlaAspAla 205
Db 588 AGATCAGCGTTTGGGCAAGTGTCAAAGAGCAAGACAAATTTATACAACTAATAAGAA 647
QY 206 ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeu 225
Db 648 GTGATACTCTTAGCAGGAGGTTTGATGTGGCTGACATATTCCCTTCACTGAAGTTCTT 707
QY 226 HisTyrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 245
Db 708 CATGTGCTCAGTGAATGAAGGTAAGATTATGAATGCACACCAATGAAGGTAGATGCCATT 767
QY 246 LeuGluGluLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 258
Db 768 GTTGAGAATGTCATCAATGAGCACAAGAAAAATCTTGGCAATTTGGGAAAACTAATGAGCG 827
QY 259 PheGluAlaAspAsnLeuMetAspValLeuLeuLeuGlnLysAsnGlyAsnValPro 278
Db 828 TTAGGAGGTGAAGATTAAATGATGTTCTTCTAAGACTTATGAATGATGAGGCGCTTCAA 887
QY 279 ValProValThrAsnGluSerLysLysLysLysLysLysLysLysLysLysLysLysSer 298
Db 888 TTTCTTATCACCACCAAGCAACATCAAAAGCTATAATTTTTTGACATGTTTCTGCGCGGACA 947
QY 299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
Db 948 GAGACTTTCATCTCAACCAATTTGTTGGGCTATGGTGAAGAAATGGTGAAGAAATCCAGCCGTA 1007
QY 319 LeuArgLysAlaGlnGluValArgGlnValPhePheGlyGluMetGlyLysValAspGlu 338
Db 1008 TTCCGGAAGCTCAAGCAGAGAGTAAAGAGAGCATTTAGAGGAAAGAAACTTTTCGATGAA 1067
QY 339 SerArgPheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHis 358
Db 1068 AATGATGTGAGGAGCTAAACTACCTAAGTTAGTAATAAAGAAAGAACTCTAAGACTTCAT 1127
```



```
QY 359 ProProVal---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyr 377
Db ||||| :|||:||||| ||||| ||||| ||||| ||||| ||||| |||||
1128 CCACCGTTCCACTTTGCTCCCAAGAGAATGTAGGGAAGACACAATATAAACGGCTAC 1187

QY 378 GluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsn 397
Db ||||| :|||:||||| ||||| ||||| ||||| ||||| ||||| |||||
1188 ACTATTCTCTGTAAGACCAACCATCATGTTTATGTTGGGCTTTGGGAAGAGATCCAAA 1247

QY 398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr 417
Db ||||| :|||:||||| ||||| ||||| ||||| ||||| ||||| |||||
1248 TATTGGAATGACGCGAAGAACTTTTATGCCAGAGAGATTTGACAGTGTCTTAAGGATTTT 1307

QY 418 LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIle 437
Db ||||| :|||:||||| ||||| ||||| ||||| ||||| ||||| |||||
1308 GTTGTAAATAATTTTGAATATCTTCCATTTGGTGGGGAAGGAGATTTGCTGGGATT 1367

QY 438 ThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrp 457
Db ||||| :|||:||||| ||||| ||||| ||||| ||||| ||||| |||||
1368 TCGTTTGGCTTAGCTAATGCTTATTGTCATTGGCTCAATTACTATATCACTTCGATTGG 1427

QY 458 GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyAla 477
Db ||||| :|||:||||| ||||| ||||| ||||| ||||| ||||| |||||
1428 AAACCTCCCTGCTGGAAATCGAACCAAGCGACTTGGACTGACTGAGTTGGTGGAGTAAC 1487

QY 478 LeuArgLysLysIleAspLeuLysLeuIleProIleProTyrGln 492
Db ||||| :|||:||||| ||||| ||||| ||||| ||||| ||||| |||||
1488 GCCGCTAGAAAAAGTGACCTTTACTTGGTTGGGACTCCTCTTATCAA 1532
```

Search completed: May 31, 2006, 09:32:49
Job time : 1756 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame plus p2n model

Run on: May 31, 2006, 08:37:04 ; Search time 30 Seconds
(without alignments)
2944.349 Million cell updates/sec

US-10-759-813-2
Title: US-10-759-813-2
Perfect score: 2574
Sequence: 1 MEQKNLSFPSSILISFLVL IKIDKLPIPYOVSLGSNIS 500

Scoring table: BLOSUM62

Searched: 246837 seqs, 58886990 residues

Total number of hits satisfying chosen parameters: 493674

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```

Command line parameters:
-MODEL=frame_pzn.model -DEV=xlh
-Q=/abs/ABSWEB/spool/US10759813/runat_31052006_060316_2859/app_query.fasta_1
-DB=Publshed Applications NA New -SUFFIX=inpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -QTWAT=1 -END=-1 -MATRIX=blosom62
-TRANS=human40 -LI=LIST-45 -DLOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abs02h
-USER=US10759813 @CGN 1 18 @runat_31052006_060316_2859 -NCPU=3
-NO MWAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

```

```
Database : Published Applications_NA_New.*
1 : /ENC_Celerra_SIDS3/prodata/2/pubnpna/US09_NEW_PUB.seq.*
2 : /ENC_Celerra_SIDS3/prodata/2/pubnpna/US06_NEW_PUB.seq.*
3 : /ENC_Celerra_SIDS3/prodata/2/pubnpna/US07_NEW_PUB.seq.*
4 : /ENC_Celerra_SIDS3/prodata/2/pubnpna/US08_NEW_PUB.seq.*
5 : /ENC_Celerra_SIDS3/prodata/2/pubnpna/FCI_NEW_PUB.seq.*
6 : /ENC_Celerra_SIDS3/prodata/2/pubnpna/US10_NEW_PUB.seq.*
7 : /ENC_Celerra_SIDS3/prodata/2/pubnpna/US11_NEW_PUB.seq.*
8 : /ENC_Celerra_SIDS3/prodata/2/pubnpna/US60_NEW_PUB.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	1099	42.7	1751	6	US-10-953-349-22681	A	Sequence 22681, A
2	1054.5	41.0	1504	6	US-10-953-349-24167	A	Sequence 24167, A
3	944	36.7	1632	6	US-10-953-349-32482	A	Sequence 32482, A
4	943.5	36.7	1766	6	US-10-953-349-32468	A	Sequence 32468, A
5	911.5	35.4	1711	6	US-10-953-349-33890	A	Sequence 33890, A
6	840	32.6	1915	6	US-10-953-349-13595	A	Sequence 13595, A
7	828.5	32.2	1737	6	US-10-953-349-14590	A	Sequence 14590, A
8	817.5	31.8	1617	6	US-10-953-349-2239	Ap	Sequence 2239, Ap
9	816.5	31.7	1586	6	US-10-953-349-4847	Ap	Sequence 4847, Ap


```
Db      306  CTTGTCAAAATGGATTCATTCAAAAGAAAGGATCCCAATCAACCTCACTGAAGCAGTACTT 365
QY      179  GlyIleIleSerIleMetIleArgThrSerValGly---AsnCyValysGlnLysGlu 197
Db      366  ACATCAATTACAAATTCCTCAAGGGCTGCTTTGGCAAGAAATTCGAAGACCAAGAA 425
QY      198  ArgLeuLeuSerValAlaAspAlaValAsnGluAlaAlaThrSerPheGlyThrAlaAsp 217
Db      426  AAATTTATATCACTGGTTAAATAAACATCAAAACTTGTCTGACAGTTTTCGCATTGAAGAT 485
QY      218  AlaPheProThrTrpLysLeuLeuHisTyrIleIleGlyAlaGluSerLysProArgArg 237
Db      486  TTGTTTCTCTCTTACTTGGCTTCAACATGTCACTGGCTTGAGGGCTAAGCTTGAGAGG 545
QY      238  LeuHisGlnGluLeuAspPheLeuGluGluLeuLeuAsnGluHisLys---AlaAsn 256
Db      546  TTGCATCAACAGCTGATCAGATAATGGAAACATCATCAATGAGCATAAAGAGGCAAAAT 605
QY      257  LysProPheGluAlaAsp-----AsnLeuMetAspValLeuLeuAsn 270
Db      606  TCGAAAGCCCAAGGATGACCAAGTAGAAGCAGAGAGAGATCTTGTGATGTTCTCATCAA 665
QY      271  LeuGlnLysAsnGlyAsnValProValProValThrAsnGluSerIleLysAlaSerVal 290
Db      666  TATGAGGATGGGAGCAAGAGGATTTTCTTAACATAGAAACAAAGATCAAGGCCATAAAT 725
QY      291  LeuGlnMetPheThrAlaGlySerGluThrThrSerLysAlaThrGluTrpValMetAla 310
Db      726  CTGGACATTTTCTCTGCTGGAGTGAGACAACAGCAACCAATAGATATGGGCAATGGCT 785
QY      311  GluLeuMetLysAsnProThrGluLeuArgLysAlaGlnGluValArgGlnValPhe 330
Db      786  GAAATGGTAAGAATCCACACAGTAATGAAGAAAGCACAACTCTGAGGTGAGAGAGGTATGC 845
QY      331  GlyLysMetClyLysValAspGluSerArgPheHisAspLeuLysPhePheLysLeuVal 350
Db      846  AATATGAAAGCAAGGGTTCATGAAATTCATCAATCAATCAATCAATCAATCAATCAAT 905
QY      351  ValLysGluThrLeuArgLeuHisProVal---ValleuLeuProArgGluCysArg 369
Db      906  GTGAAAGAGACCTTGAGGTGACACCTCCAGCTCCCTTTTGTGTCGAAGAGAAATGTGT 965
QY      370  GluThrThrArgIleAspGlyTyrGluIleHisProAsnThrArgIleValValAsnAla 389
Db      966  CAAACATGTGAGATACATGTTATCATACATACCAGCCAAACCAAGTCTATGTCATGCTC 1025
QY      390  TrpAlaIleGlyArgAspProAsnThrTrpSerGluProGlyLysPheAsnProGluArg 409
Db      1026  TGGGCAATTGGAAGAGATCCAAACTATTGACTGTAATCAGAGAGGTTTATCCTGAGAGG 1085
QY      410  PheLysAspCysAlaIleAspTyrLysGlyThrThrPheGluLeuValProPheGlyAla 429
Db      1086  TTCATTGATGACACTATTGACTACAAAGGGAGTAATTTTGGATTCAATCCTTTTGTGTCT 1145
QY      430  GlyLysArgIleCysProGlyIleThrSerAlaIleThrAsnLeuGluTyrValIleIle 449
Db      1146  GGAAGAGGATATGCCAGGAGACACATTTGCTTTGAGAGCTCCGGAATAGCCCTTGCA 1205
QY      450  AsnLeuLeuTyrHisPheAsnTrpGluLeuAlaAspGlyIleThrProGlnThrLeuAsp 469
Db      1206  ATGTTGTTGATCATTGATTTGATTTGAAAGCTTCCAAGTGAATGAGAAGTGAGAACTGGAC 1265
QY      470  MetThrGluAlaIleGlyAlaLeuArgLysLysIleAspLeuLysLeuIleProIle 489
Db      1266  ATGAGTGAGGATTTTGGAGTCAACAATAAAGAAAGATAATCTATTCTTGTGTTCTCTTT 1325
QY      490  ProTyr 491
Db      1326  CCATAT 1331
QY      491  ProTyr 491
Db      1331  CCATAT 1331
QY      492  ProTyr 492
Db      1332  CCATAT 1337
QY      493  ProTyr 493
Db      1333  CCATAT 1337
QY      494  ProTyr 494
Db      1334  CCATAT 1337
QY      495  ProTyr 495
Db      1335  CCATAT 1337
QY      496  ProTyr 496
Db      1336  CCATAT 1337
QY      497  ProTyr 497
Db      1337  CCATAT 1337
QY      498  ProTyr 498
Db      1338  CCATAT 1337
QY      499  ProTyr 499
Db      1339  CCATAT 1337
QY      500  ProTyr 500
Db      1340  CCATAT 1337
QY      501  ProTyr 501
Db      1341  CCATAT 1337
QY      502  ProTyr 502
Db      1342  CCATAT 1337
QY      503  ProTyr 503
Db      1343  CCATAT 1337
QY      504  ProTyr 504
Db      1344  CCATAT 1337
QY      505  ProTyr 505
Db      1345  CCATAT 1337
QY      506  ProTyr 506
Db      1346  CCATAT 1337
QY      507  ProTyr 507
Db      1347  CCATAT 1337
QY      508  ProTyr 508
Db      1348  CCATAT 1337
QY      509  ProTyr 509
Db      1349  CCATAT 1337
QY      510  ProTyr 510
Db      1350  CCATAT 1337
QY      511  ProTyr 511
Db      1351  CCATAT 1337
QY      512  ProTyr 512
Db      1352  CCATAT 1337
QY      513  ProTyr 513
Db      1353  CCATAT 1337
QY      514  ProTyr 514
Db      1354  CCATAT 1337
QY      515  ProTyr 515
Db      1355  CCATAT 1337
QY      516  ProTyr 516
Db      1356  CCATAT 1337
QY      517  ProTyr 517
Db      1357  CCATAT 1337
QY      518  ProTyr 518
Db      1358  CCATAT 1337
QY      519  ProTyr 519
Db      1359  CCATAT 1337
QY      520  ProTyr 520
Db      1360  CCATAT 1337
QY      521  ProTyr 521
Db      1361  CCATAT 1337
QY      522  ProTyr 522
Db      1362  CCATAT 1337
QY      523  ProTyr 523
Db      1363  CCATAT 1337
QY      524  ProTyr 524
Db      1364  CCATAT 1337
QY      525  ProTyr 525
Db      1365  CCATAT 1337
QY      526  ProTyr 526
Db      1366  CCATAT 1337
QY      527  ProTyr 527
Db      1367  CCATAT 1337
QY      528  ProTyr 528
Db      1368  CCATAT 1337
QY      529  ProTyr 529
Db      1369  CCATAT 1337
QY      530  ProTyr 530
Db      1370  CCATAT 1337
QY      531  ProTyr 531
Db      1371  CCATAT 1337
QY      532  ProTyr 532
Db      1372  CCATAT 1337
QY      533  ProTyr 533
Db      1373  CCATAT 1337
QY      534  ProTyr 534
Db      1374  CCATAT 1337
QY      535  ProTyr 535
Db      1375  CCATAT 1337
QY      536  ProTyr 536
Db      1376  CCATAT 1337
QY      537  ProTyr 537
Db      1377  CCATAT 1337
QY      538  ProTyr 538
Db      1378  CCATAT 1337
QY      539  ProTyr 539
Db      1379  CCATAT 1337
QY      540  ProTyr 540
Db      1380  CCATAT 1337
QY      541  ProTyr 541
Db      1381  CCATAT 1337
QY      542  ProTyr 542
Db      1382  CCATAT 1337
QY      543  ProTyr 543
Db      1383  CCATAT 1337
QY      544  ProTyr 544
Db      1384  CCATAT 1337
QY      545  ProTyr 545
Db      1385  CCATAT 1337
QY      546  ProTyr 546
Db      1386  CCATAT 1337
QY      547  ProTyr 547
Db      1387  CCATAT 1337
QY      548  ProTyr 548
Db      1388  CCATAT 1337
QY      549  ProTyr 549
Db      1389  CCATAT 1337
QY      550  ProTyr 550
Db      1390  CCATAT 1337
QY      551  ProTyr 551
Db      1391  CCATAT 1337
QY      552  ProTyr 552
Db      1392  CCATAT 1337
QY      553  ProTyr 553
Db      1393  CCATAT 1337
QY      554  ProTyr 554
Db      1394  CCATAT 1337
QY      555  ProTyr 555
Db      1395  CCATAT 1337
QY      556  ProTyr 556
Db      1396  CCATAT 1337
QY      557  ProTyr 557
Db      1397  CCATAT 1337
QY      558  ProTyr 558
Db      1398  CCATAT 1337
QY      559  ProTyr 559
Db      1399  CCATAT 1337
QY      560  ProTyr 560
Db      1400  CCATAT 1337
QY      561  ProTyr 561
Db      1401  CCATAT 1337
QY      562  ProTyr 562
Db      1402  CCATAT 1337
QY      563  ProTyr 563
Db      1403  CCATAT 1337
QY      564  ProTyr 564
Db      1404  CCATAT 1337
QY      565  ProTyr 565
Db      1405  CCATAT 1337
QY      566  ProTyr 566
Db      1406  CCATAT 1337
QY      567  ProTyr 567
Db      1407  CCATAT 1337
QY      568  ProTyr 568
Db      1408  CCATAT 1337
QY      569  ProTyr 569
Db      1409  CCATAT 1337
QY      570  ProTyr 570
Db      1410  CCATAT 1337
QY      571  ProTyr 571
Db      1411  CCATAT 1337
QY      572  ProTyr 572
Db      1412  CCATAT 1337
QY      573  ProTyr 573
Db      1413  CCATAT 1337
QY      574  ProTyr 574
Db      1414  CCATAT 1337
QY      575  ProTyr 575
Db      1415  CCATAT 1337
QY      576  ProTyr 576
Db      1416  CCATAT 1337
QY      577  ProTyr 577
Db      1417  CCATAT 1337
QY      578  ProTyr 578
Db      1418  CCATAT 1337
QY      579  ProTyr 579
Db      1419  CCATAT 1337
QY      580  ProTyr 580
Db      1420  CCATAT 1337
QY      581  ProTyr 581
Db      1421  CCATAT 1337
QY      582  ProTyr 582
Db      1422  CCATAT 1337
QY      583  ProTyr 583
Db      1423  CCATAT 1337
QY      584  ProTyr 584
Db      1424  CCATAT 1337
QY      585  ProTyr 585
Db      1425  CCATAT 1337
QY      586  ProTyr 586
Db      1426  CCATAT 1337
QY      587  ProTyr 587
Db      1427  CCATAT 1337
QY      588  ProTyr 588
Db      1428  CCATAT 1337
QY      589  ProTyr 589
Db      1429  CCATAT 1337
QY      590  ProTyr 590
Db      1430  CCATAT 1337
QY      591  ProTyr 591
Db      1431  CCATAT 1337
QY      592  ProTyr 592
Db      1432  CCATAT 1337
QY      593  ProTyr 593
Db      1433  CCATAT 1337
QY      594  ProTyr 594
Db      1434  CCATAT 1337
QY      595  ProTyr 595
Db      1435  CCATAT 1337
QY      596  ProTyr 596
Db      1436  CCATAT 1337
QY      597  ProTyr 597
Db      1437  CCATAT 1337
QY      598  ProTyr 598
Db      1438  CCATAT 1337
QY      599  ProTyr 599
Db      1439  CCATAT 1337
QY      600  ProTyr 600
Db      1440  CCATAT 1337
QY      601  ProTyr 601
Db      1441  CCATAT 1337
QY      602  ProTyr 602
Db      1442  CCATAT 1337
QY      603  ProTyr 603
Db      1443  CCATAT 1337
QY      604  ProTyr 604
Db      1444  CCATAT 1337
QY      605  ProTyr 605
Db      1445  CCATAT 1337
QY      606  ProTyr 606
Db      1446  CCATAT 1337
QY      607  ProTyr 607
Db      1447  CCATAT 1337
QY      608  ProTyr 608
Db      1448  CCATAT 1337
QY      609  ProTyr 609
Db      1449  CCATAT 1337
QY      610  ProTyr 610
Db      1450  CCATAT 1337
QY      611  ProTyr 611
Db      1451  CCATAT 1337
QY      612  ProTyr 612
Db      1452  CCATAT 1337
QY      613  ProTyr 613
Db      1453  CCATAT 1337
QY      614  ProTyr 614
Db      1454  CCATAT 1337
QY      615  ProTyr 615
Db      1455  CCATAT 1337
QY      616  ProTyr 616
Db      1456  CCATAT 1337
QY      617  ProTyr 617
Db      1457  CCATAT 1337
QY      618  ProTyr 618
Db      1458  CCATAT 1337
QY      619  ProTyr 619
Db      1459  CCATAT 1337
QY      620  ProTyr 620
Db      1460  CCATAT 1337
QY      621  ProTyr 621
Db      1461  CCATAT 1337
QY      622  ProTyr 622
Db      1462  CCATAT 1337
QY      623  ProTyr 623
Db      1463  CCATAT 1337
QY      624  ProTyr 624
Db      1464  CCATAT 1337
QY      625  ProTyr 625
Db      1465  CCATAT 1337
QY      626  ProTyr 626
Db      1466  CCATAT 1337
QY      627  ProTyr 627
Db      1467  CCATAT 1337
QY      628  ProTyr 628
Db      1468  CCATAT 1337
QY      629  ProTyr 629
Db      1469  CCATAT 1337
QY      630  ProTyr 630
Db      1470  CCATAT 1337
QY      631  ProTyr 631
Db      1471  CCATAT 1337
QY      632  ProTyr 632
Db      1472  CCATAT 1337
QY      633  ProTyr 633
Db      1473  CCATAT 1337
QY      634  ProTyr 634
Db      1474  CCATAT 1337
QY      635  ProTyr 635
Db      1475  CCATAT 1337
QY      636  ProTyr 636
Db      1476  CCATAT 1337
QY      637  ProTyr 637
Db      1477  CCATAT 1337
QY      638  ProTyr 638
Db      1478  CCATAT 1337
QY      639  ProTyr 639
Db      1479  CCATAT 1337
QY      640  ProTyr 640
Db      1480  CCATAT 1337
QY      641  ProTyr 641
Db      1481  CCATAT 1337
QY      642  ProTyr 642
Db      1482  CCATAT 1337
QY      643  ProTyr 643
Db      1483  CCATAT 1337
QY      644  ProTyr 644
Db      1484  CCATAT 1337
QY      645  ProTyr 645
Db      1485  CCATAT 1337
QY      646  ProTyr 646
Db      1486  CCATAT 1337
QY      647  ProTyr 647
Db      1487  CCATAT 1337
QY      648  ProTyr 648
Db      1488  CCATAT 1337
QY      649  ProTyr 649
Db      1489  CCATAT 1337
QY      650  ProTyr 650
Db      1490  CCATAT 1337
QY      651  ProTyr 651
Db      1491  CCATAT 1337
QY      652  ProTyr 652
Db      1492  CCATAT 1337
QY      653  ProTyr 653
Db      1493  CCATAT 1337
QY      654  ProTyr 654
Db      1494  CCATAT 1337
QY      655  ProTyr 655
Db      1495  CCATAT 1337
QY      656  ProTyr 656
Db      1496  CCATAT 1337
QY      657  ProTyr 657
Db      1497  CCATAT 1337
QY      658  ProTyr 658
Db      1498  CCATAT 1337
QY      659  ProTyr 659
Db      1499  CCATAT 1337
QY      660  ProTyr 660
Db      1500  CCATAT 1337
QY      661  ProTyr 661
Db      1501  CCATAT 1337
QY      662  ProTyr 662
Db      1502  CCATAT 1337
QY      663  ProTyr 663
Db      1503  CCATAT 1337
QY      664  ProTyr 664
Db      1504  CCATAT 1337
QY      665  ProTyr 665
Db      1505  CCATAT 1337
QY      666  ProTyr 666
Db      1506  CCATAT 1337
QY      667  ProTyr 667
Db      1507  CCATAT 1337
QY      668  ProTyr 668
Db      1508  CCATAT 1337
QY      669  ProTyr 669
Db      1509  CCATAT 1337
QY      670  ProTyr 670
Db      1510  CCATAT 1337
QY      671  ProTyr 671
Db      1511  CCATAT 1337
QY      672  ProTyr 672
Db      1512  CCATAT 1337
QY      673  ProTyr 673
Db      1513  CCATAT 1337
QY      674  ProTyr 674
Db      1514  CCATAT 1337
QY      675  ProTyr 675
Db      1515  CCATAT 1337
QY      676  ProTyr 676
Db      1516  CCATAT 1337
QY      677  ProTyr 677
Db      1517  CCATAT 1337
QY      678  ProTyr 678
Db      1518  CCATAT 1337
QY      679  ProTyr 679
Db      1519  CCATAT 1337
QY      680  ProTyr 680
Db      1520  CCATAT 1337
QY      681  ProTyr 681
Db      1521  CCATAT 1337
QY      682  ProTyr 682
Db      1522  CCATAT 1337
QY      683  ProTyr 683
Db      1523  CCATAT 1337
QY      684  ProTyr 684
Db      1524  CCATAT 1337
QY      685  ProTyr 685
Db      1525  CCATAT 1337
QY      686  ProTyr 686
Db      1526  CCATAT 1337
QY      687  ProTyr 687
Db      1527  CCATAT 1337
QY      688  ProTyr 688
Db      1528  CCATAT 1337
QY      689  ProTyr 689
Db      1529  CCATAT 1337
QY      690  ProTyr 690
Db      1530  CCATAT 1337
QY      691  ProTyr 691
Db      1531  CCATAT 1337
QY      692  ProTyr 692
Db      1532  CCATAT 1337
QY      693  ProTyr 693
Db      1533  CCATAT 1337
QY      694  ProTyr 694
Db      1534  CCATAT 1337
QY      695  ProTyr 695
Db      1535  CCATAT 1337
QY      696  ProTyr 696
Db      1536  CCATAT 1337
QY      697  ProTyr 697
Db      1537  CCATAT 1337
QY      698  ProTyr 698
Db      1538  CCATAT 1337
QY      699  ProTyr 699
Db      1539  CCATAT 1337
QY      700  ProTyr 700
Db      1540  CCATAT 1337
QY      701  ProTyr 701
Db      1541  CCATAT 1337
QY      702  ProTyr 702
Db      1542  CCATAT 1337
QY      703  ProTyr 703
Db      1543  CCATAT 1337
QY      704  ProTyr 704
Db      1544  CCATAT 1337
QY      705  ProTyr 705
Db      1545  CCATAT 1337
QY      706  ProTyr 706
Db      1546  CCATAT 1337
QY      707  ProTyr 707
Db      1547  CCATAT 1337
QY      708  ProTyr 708
Db      1548  CCATAT 1337
QY      709  ProTyr 709
Db      1549  CCATAT 1337
QY      710  ProTyr 710
Db      1550  CCATAT 1337
QY      711  ProTyr 711
Db      1551  CCATAT 1337
QY      712  ProTyr 712
Db      1552  CCATAT 1337
QY      713  ProTyr 713
Db      1553  CCATAT 1337
QY      714  ProTyr 714
Db      1554  CCATAT 1337
QY      715  ProTyr 715
Db      1555  CCATAT 1337
QY      716  ProTyr 716
Db      1556  CCATAT 1337
QY      717  ProTyr 717
Db      1557  CCATAT 1337
QY      718  ProTyr 718
Db      1558  CCATAT 1337
QY      719  ProTyr 719
Db      1559  CCATAT 1337
QY      720  ProTyr 720
Db      1560  CCATAT 1337
QY      721  ProTyr 721
Db      1561  CCATAT 1337
QY      722  ProTyr 722
Db      1562  CCATAT 1337
QY      723  ProTyr 723
Db      1563  CCATAT 1337
QY      724  ProTyr 724
Db      1564  CCATAT 1337
QY      725  ProTyr 725
Db      1565  CCATAT 1337
QY      726  ProTyr 726
Db      1566  CCATAT 1337
QY      727  ProTyr 727
Db      1567  CCATAT 1337
QY      728  ProTyr 728
Db      1568  CCATAT 1337
QY      729  ProTyr 729
Db      1569  CCATAT 1337
QY      730  ProTyr 730
Db      1570  CCATAT 1337
QY      731  ProTyr 731
Db      1571  CCATAT 1337
QY      732  ProTyr 732
Db      1572  CCATAT 1337
QY      733  ProTyr 733
Db      1573  CCATAT 1337
QY      734  ProTyr 734
Db      1574  CCATAT 1337
QY      735  ProTyr 735
Db      1575  CCATAT 1337
QY      736  ProTyr 736
Db      1576  CCATAT 1337
QY      737  ProTyr 737
Db      1577  CCATAT 1337
QY      738  ProTyr 738
Db      1578  CCATAT 1337
QY      739  ProTyr 739
Db      1579  CCATAT 1337
QY      740  ProTyr 740
Db      1580  CCATAT 1337
QY      741  ProTyr 741
Db      1581  CCATAT 1337
QY      742  ProTyr 742
Db      1582  CCATAT 1337
QY      743  ProTyr 743
Db      1583  CCATAT 1337
QY      744  ProTyr 744
Db      1584  CCATAT 1337
QY      745  ProTyr 745
Db      1585  CCATAT 1337
QY      746  ProTyr 746
Db      1586  CCATAT 1337
QY      747  ProTyr 747
Db      1587  CCATAT 1337
QY      748  ProTyr 748
Db      1588  CCATAT 1337
QY      749  ProTyr 749
Db      1589  CCATAT 1337
QY      750  ProTyr 750
Db      1590  CCATAT 1337
QY      751  ProTyr 751
Db      1591  CCATAT 1337
QY      752  ProTyr 752
Db      1592  CCATAT 1337
QY      753  ProTyr 753
Db      1593  CCATAT 1337
QY      754  ProTyr 754
Db      1594  CCATAT 1337
QY      755  ProTyr 755
Db      1595  CCATAT 1337
QY      756  ProTyr 756
Db      1596  CCATAT 1337
QY      757  ProTyr 757
Db      1597  CCATAT 1337
QY      758  ProTyr 758
Db      1598  CCATAT 1337
QY      759  ProTyr 759
Db      1599  CCATAT 1337
QY      760  ProTyr 760
Db      1600  CCATAT 1337
QY      761  ProTyr 761
Db      1601  CCATAT 1337
QY      762  ProTyr 762
Db      1602  CCATAT 1337
QY      763  ProTyr 763
Db      1603  CCATAT 1337
QY      764  ProTyr 764
Db      1604  CCATAT 1337
QY      765  ProTyr 765
Db      1605  CCATAT 1337
QY      766  ProTyr 766
Db      1606  CCATAT 1337
QY      767  ProTyr 767
Db      1607  CCATAT 1337
QY      768  ProTyr 768
Db      1608  CCATAT 1337
QY      769  ProTyr 769
Db      1609  CCATAT 1337
QY      770  ProTyr 770
Db      1610  CCATAT 1337
QY      771  ProTyr 771
Db      1611  CCATAT 1337
QY      772  ProTyr 772
Db      1612  CCATAT 1337
QY      773  ProTyr 773
Db      1613  CCATAT 1337
QY      774  ProTyr 774
Db      1614  CCATAT 1337
QY      775  ProTyr 775
Db      1615  CCATAT 1337
QY      776  ProTyr 776
Db      1616  CCATAT 1337
QY      777  ProTyr 777
Db      1617  CCATAT 1337
QY      778  ProTyr 778
Db      1618  CCATAT 1337
QY      779  ProTyr 779
Db      1619  CCATAT 1337
QY      780  ProTyr 780
Db      1620  CCATAT 1337
QY      781  ProTyr 781
Db      1621  CCATAT 1337
QY      782  ProTyr 782
Db      1622  CCATAT 1337
QY      783  ProTyr 783
Db      1623  CCATAT 1337
QY      784  ProTyr 784
Db      1624  CCATAT 1337
QY      785  ProTyr 785
Db      1625  CCATAT 1337
QY      786  ProTyr 786
Db      1626  CCATAT 1337
QY      787  ProTyr 787
Db      1627  CCATAT 1337
QY      788  ProTyr 788
Db      1628  CCATAT 1337
QY      789  ProTyr 789
Db      1629  CCATAT 1337
QY      790  ProTyr 790
Db      1630  CCATAT 1337
QY      791  ProTyr 791
Db      1631  CCATAT 1337
QY      792  ProTyr 792
Db      1632  CCATAT 1337
QY      793  ProTyr 793
Db      1633  CCATAT 1337
QY      794  ProTyr 794
Db      1634  CCATAT 1337
QY      795  ProTyr 795
Db      1635  CCATAT 1337
QY      796  ProTyr 796
Db      1636  CCATAT 1337
QY      797  ProTyr 797
Db      1637  CCATAT 1337
QY      798  ProTyr 798
Db      1638  CCATAT 1337
QY      799  ProTyr 799
Db      1639  CCATAT 1337
QY      800  ProTyr 800
Db      1640  CCATAT 1337
QY      801  ProTyr 801
Db      1641  CCATAT 1337
QY      802  ProTyr 802
Db      1642  CCATAT 1337
QY      803  ProTyr 803
Db      1643  CCATAT 1337
QY      804  ProTyr 804
Db      1644  CCATAT 1337
QY      805  ProTyr 805
Db      1645  CCATAT 1337
QY      806  ProTyr 806
Db      1646  CCATAT 1337
QY      807  ProTyr 807
Db      1647  CCATAT 1337
QY      808  ProTyr 808
Db      1648  CCATAT 1337
QY      809  ProTyr 809
Db      1649  CCATAT 1337
QY      810  ProTyr 810
Db      1650  CCATAT 1337
QY      811  ProTyr 811
Db      1651  CCATAT 1337
QY      812  ProTyr 812
Db      1652  CCATAT 1337
QY      813  ProTyr 813
Db      1653  CCATAT 1337
QY      814  ProTyr 814
Db      1654  CCATAT 1337
QY      815  ProTyr 815
Db      1655  CCATAT 1337
QY      816  Pro
```

```
Db 689 GACATTTTCCCTCTCGCGTCTCGTGGCTCCTTAGCAGCGTCCCGCGCAGGATACAG 748
Qy 237 ArgLeuHisGlnGluIleAspAspLeuLeuGluGluIleLeuAsnGluHisLysAlaAsn 256
Db 749 CGCCATAGCAACATGAAGCTGTTTCATGGACACCATCATCAAGAGCACCAGGTGAAC 808
Qy 257 LysProPheGluAlaAsp-----AsnLeuMetAspValLeuLeuAsnLeu 271
Db 809 AGCGGCCCGCAGTGGCGCGCGACAAAGAAAGAGACTTGTCTGACGCTGCTCTGAGATC 868
Qy 272 GlnLysAsnGlyAsnValProValProValThrAsnGluSerIleLysAlaSerValLeu 291
Db 869 CAAAGGAGGAGGACTCGCAGTATCACTACCACTGACAAATCAAGACCGTCACTG 928
Qy 292 GlnMetPheThrAlaGlySerGluThrThrSerLysAlaThrGluTrpValMetAlaGlu 311
Db 929 GACATGTTTGGCGCGCGCAGCAGAGCTGCGCGACGACGCTGCTGCGCGATGCGGAG 988
Qy 312 LeuMetLysAsnProThrGluLeuArgLysAlaGlnGluValArgGlnValPheGly 331
Db 989 CTGATACGGAACCCCGGGTTATGCGGAAGCGCGCAAGACGAGGTCCGCGCAACTCGCC 1048
Qy 332 GluMetGlyLysValAspGluSerArgPheHisAspLeuLysPhePheLysLeuVal 351
Db 1049 GGGCAGCAGAGTGAGGAGGAGCCACCTGGCAGATCTGCGATCTTGGGTTTGTATC 1108
Qy 352 LysGluThrLeuArgLysHisProProVal---ValLeuIleProArgGluCysArgGlu 370
Db 1109 AAGGAGAGCTGAGGATGACCCCGCGCGCCATTGCTGCTGCCCGTAGGTGCGGAGC 1168
Qy 371 ThrThrArgIleAspGlyTyroGluLeHisProAsnThrArgIleValAlaAsnAlaTrp 390
Db 1169 CGGTGTCAAGTTCTCGGCTGAGCGTGGCGGAGGGGTGATGATGATGATGATGATG 1228
Qy 391 AlaIleGlyArgAspProAsnThrTrpSerGluProGlyLysPheAsnProGluArgPhe 410
Db 1229 GCGATCGGATGACCGCGGCACTGGGAGGACCGGAGGAGTTCGACCGGAGGGTTC 1288
Qy 411 LysAspCysAlaIleAspTyroLysGlyThrThrPheGluLeuValProPheGlyAlaGly 430
Db 1289 GAGCAGAACGGAGGAGACTTCAAGGGGCGGACTTCGAGTTCTGCTGCCGTTCGGCGCGG 1348
Qy 431 LysArgIleCysProGlyThrSerAlaIleThrAsnLeuGluTyroValIleIleAsn 450
Db 1349 AGGAGATATGCCCCGATGCGCTTTCGGGCTGGCGCACGTCGAGCTCGCGCTCGCGCG 1408
Qy 451 LeuLeuTyroHisPheAsnTrpGluLeuAlaAspGlyIleThrProGlnThrLeuAspMet 470
Db 1409 CTGCTGTTCCACTTGACTGGAGGTGCCAGCGGGGCGCGCGGCGGAGGATCTGACATG 1468
Qy 471 ThrGluAlaIleGlyClyAlaLeuArgLysLysIleAspLeuLysLeuIleProIle 489
Db 1469 ACCGAGGAGTTTGGCGTCACGCGACCGGCTCGGCTCGACCTTGTGCTGCTGCGCGTG 1525

RESULT 4
US-10-759-349-32468
; Sequence 32468, Application US/109533349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 32468
; LENGTH: 1766
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-759-349-32468
```

```
Alignment Scores:
Pred. No.: 8.76e-92 Length: 1766
Score: 943.50 Matches: 190
Percent Similarity: 60.9% Conservative: 115
Best Local Similarity: 37.9% Mismatches: 177
Query Match: 36.7% Indels: 19
DB: 6 Gaps: 9

US-10-759-813-2 (1-500) x US-10-953-349-32468 (1-1766)

Qy 7 SerPheProSerIleLeuIleSerPheLeuLeuValLeuIleLeuValValVal----- 24
Db 83 ACCATTATCAGGTCGCTCGCGTAGCTCTGTTTCGCTCATCATGATTGTCGGGCG 142
Qy 25 MetArgLeuTrpLysLysGlnAsnPro-----ProProGlyProTrpLysPhe 40
Db 143 CGCGCGCGCTGGGGCGCGTGGCGACGCGCTGCGTATGGCGCTTGGCGCGTGGCAGCTG 202
Qy 41 ProIleIleGlyAsnLeuProHisLeuLeuLeuThrSerAspLeuLeuGlyHisGluArgPhe 60
Db 203 CCGGTGATCGCAGACCTGCACACCTCTGCTGCGCGGCGAGCTCCCTCACCGCGCGATG 262
Qy 61 ArgAlaLeuAlaGlnIleTyroGlyProValMetSerLeuGlnIleGlyGlnValSerAla 80
Db 263 CGCGACCTGGCGCGCAGCTACGCGCGCGCTGCTGCTCCAGCTCGGCCAGGTGAAGACG 322
Qy 81 ValValIleSerSerAlaGluAlaLysGluValMetLysThrGlnAlaAspAlaPhe 100
Db 323 CTGTGTGTCTCTCTCGGAGGGGCGCGGAGGTGATGAAGACACGACACCATGTTTC 382
Qy 101 AlaGlnArgProIleValLeuAspAlaGlnIleValPheTyroAsnArgLysAspValLeu 120
Db 383 GCCACGCGCGCTGAGCACCCACCATGCGGTGCTCTCTACCGCGCGCAGGACATCGTC 442
Qy 121 PheAlaSerTyroGlyAspHisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeu 140
Db 443 TTCGGCCCTACGAGAACTACTGGCGCAGCTCCGAGATCGCGTGTCCGAGTCTTC 502
Qy 141 SerAlaLysLysValGlnSerSerArgLeuIleArgGluGluGluMetGluAspAlaIle 160
Db 503 ACCGCGCGCGCTCTCTCTCGGCCCATCGCGAGGAGGAGGTCCGCCACGCGCTC 562
Qy 161 ThrPheLeu-----ArgSerLysAlaGlySerProValAsnIleThrLysIleIleTyro 178
Db 563 CGGTGTGTCGCGGAGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTATCC 622
Qy 179 GlyIleIleIleSerIleMetIleArgThrSerValGlyAsn---CysLysGlnLysGlu 197
Db 623 ACCTGTCGTCAGCAGCAGCAGCGCGCGCTCATAGCGCAGCGGTGCGAGGAGCGCAG 682
Qy 198 ArgLeuLeuSerValAlaAspAlaValAsnGluAlaThrSerPheGlyThrAlaAsp 217
Db 683 CGGTCTCTCCGGAGCTCGACCGCATCTGTCAGCTCGCGTGGGGTTTCAACCTGCGCGAC 742
Qy 218 AlaPheProThrTrpLysLeuHisTyroIleIleGlyAlaGluSerLysProArgArg 237
Db 743 ATGTGGCGCTCGTCGACCTTGGCGCGTGGCTCTAGTGGAGCGCGAGAGTCCGCCACACC 802
Qy 238 LeuHisGlnGluIleAspAspIleLeuGluGluIleLeuAsnGluHisLysAlaAsnLys 257
Db 803 CTGTACACCATGCTCGAGCGCATCTGTCAGGAGCACCTG---GAGAGGACGCGACGCGCG 859
Qy 258 ProPheGluAlaAspAsnLeuMetAspValLeuLeuLeuLeuGlnLysAsnGlyAsnVal 277
Db 860 GGAGGCCACGCGGAGACTTGTCTGACGCTGCTCAAGATCCAGAAAGAGGGGTGCTC 919
Qy 278 ProValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGly 297
Db 920 AAGTTTCCCATCCACATGGACGCGCTCAAGCCCATCATCTTGGAGCTATTCTCTCCCGCG 979
Qy 298 SerGluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThr 317
Db 980 AGTGAACAACAATACGACAATTGAGTGGGCCCATATATCAGAGCTGATCAACAACCCGATG 1039
```

```
QY 318 GluLeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAsp 337
Db 1040 GCCATGCAGAGCGGACAGATAGGTGGACAGCTTCCATCTAGTGAACCGGTGCC 1099
QY 338 GluSerArgPheHisAspLeuLysPhePheLysLeuValLysGluThrLeuArgLeu 357
Db 1100 GAGCAGCGCCCTAAGCAGCTCCCATACCTCGCTTGTGTCATCCGAGAGAGCTGCGGCTA 1159
QY 358 HisProValValLeuLeuProArgGluCysArgGluThrThrArgLysAspGlyTyr 377
Db 1160 CACCGCCCTGCGGTGTGTTCCGCGAGTGCAGGAACCGTGCACAGTGCAGGGATAC 1219
QY 378 GluLeHisProAsnThrArgLysValValAsnAlaTyrAlaLysValAspProAsn 397
Db 1220 GAGTGCAGCGGGGACGCGAGGTGTGTTCAATGCTTGGCGGTAGCGCGGACGAGCGC 1279
QY 398 ThrTyrSerGlu---ProGlyLysPheAsnProGluArgPheLysAspCysAla----- 414
Db 1280 TATTGCGCGGACGCGCTGAGGAGTTCCGCGCGGAACGGTTCGAGGAAGACGCAAG 1339
QY 415 IleAspTyrLysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgLysCys 434
Db 1340 GAGACTTTTGGGGTGGTGTGCTTGGCTTCCCGTGTGGCGCGCGCGGAGGATGTC 1399
QY 435 ProGlyIleThrSerAlaLysThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHis 454
Db 1400 CTGGGATGGCTTGGCTCGCGCGCGTGGAGTCCCGCTTGCAGCATGTCTTCCAC 1459
QY 455 PheAsnTyrGluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIle 474
Db 1460 TTTGACTGGAAG---CCACCGAGCGCGGAGTCTCGCGGAGTTCGACATGACAGAGCGTTC 1516
QY 475 GlyGlyAlaLeuArgLysLysIleAspLeuLysLeu-----IleProIle 489
Db 1517 GGCTCACCGCAGCGCGGACAGACGACCTCTGTGGCGCTGTCTTGGCGCTGTCTTGGCGTACCTATT 1576
QY 490 Pro 490
Db 1577 CCC 1579

RESULT 5
US-10-953-349-33880
; Sequence 33880, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 33880
; LENGTH: 1711
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
US-10-953-349-33880

Alignment Scores:
Pred. No.: 2,24e-88 Length: 1711
Score: 911.50 Matches: 191
Percent Similarity: 59.4% Conservative: 105
Best Local Similarity: 38.4% Mismatches: 180
Query Match: 35.4% Indels: 22
DB: 6 Gaps: 8

US-10-759-813-2 (1-500) x US-10-953-349-33880 (1-1711)

QY 9 ProSerIleLeuLeuSerPheLeuValLeuValLeuValValValMetArgLeuTyr 28
Db 94 CCCAGCGTCTG-----CTTCTCGTACCTTTCTCGGCATCCCTCTACTTCTTC 144
```

```
QY 29 -----LysLysGlnAsnPro-----ProProGlyProTyrLys 39
Db 145 TTGGCCACACAGAGCTAGGACGCCCGGGGGCGCGGCTCCCGCGGGCGCTGGCGC 204
QY 40 PheProIleIleGlyAsnLeuProHisLeuLeuThrSerAspLeuGlyHisGluArg 59
Db 205 CTGCGCGTGTGGGACACCTGCATCAC-----CTAGCCGGGGCGCTCCCGCACCGGTC 258
QY 60 PheArgAlaLeuAlaGlnIleTyrGlyProValMetSerLeuGlnIleGlyGlnValSer 79
Db 259 ATGCGCGACCTGCGAGGCGCCACGCGCGCTCATGTCTCGGTTCGCGAGGTCCCC 318
QY 80 AlaValValIleSerSerAlaGluAlaLysGluValMetLysThrGlnAlaAspAla 99
Db 319 GTGTGTGTGGCTCTCTCGCGCGCGCGCGGCGGATGATCGGACCCACGAGCGCGC 378
QY 100 PheAlaGlnArgProIleValLeuAspAlaGlnIleValPheTyrAsnArgLysAspVal 119
Db 379 TTGCGGTCTCGGCGCCATCGTCCCGTCTCGGCTCTGGTTCCAGGCGCGGAGGCATC 438
QY 120 LeuPheAlaSerTyrGlyAspHisTyrArgGlnMetLysLysIleTyrIleLeuGluPhe 139
Db 439 CTCCTTCGCGCCCTACGCGGACGACTGCGCGCCACTCCGCGAGGTCTGCACCCAGGAGCTC 498
QY 140 LeuSerAlaLysLysValGlnSerSerArgLeuIleArgGluGluGluMetGluAspAla 159
Db 499 CTCACCGCGCGCGCGCTCAATCTTCCGCGCGCGCGGAGGACGAGCTCCCGCGCGCTG 558
QY 160 IleThrPheLeuArgSerLysAlaGlySerProValAsnIleThrLysIleIleTyrGly 179
Db 559 CTCGCTCGCTCGCTCGCTCGAGCTCGGG---CCGCTCAACTGACCGAGAGATATCGACC 615
QY 180 IleIleIleSerIleMetIleArgThrSerValGlyAsnCysLysGlnLysGluArg--- 198
Db 616 TACATCGCGCAGCTCCACGCTGCGCGCATCATCGCGCGCGCGAGCTCAAGGACCGCGAC 675
QY 199 ---LeuLeuSerValAlaAspAlaValAsnGluAlaAlaThrSerPheGlyThrAlaAsp 217
Db 676 GGTACCTCGGATGCTCAAGGCGCTCTTCGGCATCATCCCGGGATGAGCCTCGCGGAC 735
QY 218 AlaPheProThrTyrLysLeuLeuHisTyrIleIleGlyAlaGluSerLysProArgArg 237
Db 736 CTCCTTCGCGCTCGCGCTCGCATCTCTCTCAGCGCGCTCTCTCGCGGATCCAGGCC 795
QY 238 LeuHisGlnGluIleAspAsp-IleLeuGluGluIleLeuAsnGluHisLys----- 254
Db 796 TACCTCGCGCATGCGCGACGATCATGGACGCGCATCTCCAGGACACCGGACAGAGC 855
QY 255 -AlaAsnLysProPheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAs 274
Db 856 AGCGCTGCGGAGCGGCGGAGAGGAGCTTCGTTGACGTGCTCTCTCAGACTTCAGAGGA 915
QY 274 nGlyAsnValProValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPh 294
Db 916 AGTGGACTCCCGATTTCCCGCTCACAACCTGAGAACATCAAAACAGTCTGCTGAGCATCTT 975
QY 294 eThrAlaGlySerGluThrThrSerLysAlaThrGluThrValMetAlaGluLeuMetLys 314
Db 976 TGGCGCGAGCAGCGAGACGTGCGACGACGCTGATTGGCGGATGCGCGAGCTCTCGCG 1035
QY 314 sAsnProThrGluLeuLeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetG 334
Db 1036 GAACCCCGAGGTGATGGAGAGCGCGAGGTGAGGTCCCGCAAGCGCTTTCGCGCCACGG 1095
QY 334 LysValAspGluSerArgPheHisAspLeuLysPhePheLysLeuValLysGluThr 354
Db 1096 CGCGGTGACCGAGGACAGACTAGCGCGCTGCGCTTACCTGCTGCTCATCAAGGAGTC 1155
QY 354 rLeuArgLeuHisProPro---ValValLeuIleProArgGluCysArgGluThrThrAr 373
Db 1156 GCTGCGGTGTCACCGCGCGGACGATGCTGTTCCACCGCATGTCGCGAGCGCGTGC 1215
```



```
Qy 373 gileaspGlyTyrGluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIleG1 393
Db 1216 GGTGTCGGGTACGACGTCCCGCGCGGCATCACGGTGATCGTGAAACGGGTGGCGCATCGG 1275
Qy 393 yArgAspProAsnThrTrpSerGluProGlyIlePheAsnProGluArgPheLysAspCy 413
Db 1276 CAGGACCCGCGCCACTGGACGACCCGACAAAGTTCTTCGCGAAGAGTTTCGACAGAG 1335
Qy 413 aAlaIleAspTyrLysGlyThrPheGluLeuValProPheGlyAlaGlyLysArgI1 433
Db 1336 TACCAGGACITTCAGAGGGCGAGACTTCGAGTTTCATCGGTTTCGCGGCGCGCAT 1395
Qy 433 eCyAsProGlyIleThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuTyr 453
Db 1396 CTGCCCCGCGCATAGCTTCGGCTCGCTCATATCGAGATCGCGCTCGCGCGTTCGCTGT 1455
Qy 453 rHisPheAsnTrpGluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAl 473
Db 1456 CCACITTCGACTGAGTTTCGCGGAGGGCTACGCGGCGGAGGCTGGACATGACCGAGGC 1515
Qy 473 alleGlyGlyAlaLeuArgLysIleAspLeuLysLeuIleProIlePro 490
Db 1516 ATTTGGGATTGTACGCCACGCGTTCCGACCTTTTGGTGGTGGCAACCCCT 1567
```

RESULT 6

```
US-10-953-349-13595
; Sequence 13595, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 13595
; LENGTH: 1915
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-13595
```

```
Alignment Scores:
Pred. No.: 1.21e-80 Length: 1915
Score: 840.00 Matches: 181
Percent Similarity: 56.7% Conservative: 101
Best Local Similarity: 36.4% Mismatches: 195
Query Match: 32.6% Indels: 20
DB: Gaps: 9
```

US-10-759-813-2 (1-500) x US-10-953-349-13595 (1-1915)

```
Qy 11 IleLeuIleSerPheLeuLeuValLeuIleLeuValValMetArgLeuTrpLysLys 30
Db 64 CTAACAGTTTCTTCGTCATCTTCCTATATCTTCAGCTAATAAGAGA---AACAAA 120
Qy 31 GlnAsnProProGlyProTrpLysPheProIleIleGlyAsnLeuProHisLeuLeu 50
Db 121 TACAATCTCCACCATCCCCACCAAGATACCATAATCGGCAATCTTCACCAGCTA--- 177
Qy 51 LeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIleTyrGlyProVal 70
Db 178 -----GGCACATGCCACACCGCTCTTTTCATGCACTCTCACAAATATATGCCCTCTC 231
Qy 71 MetSerLeuGlnIleGlnValSerAlaValIleSerSerAlaGlnAlaLys 90
Db 232 ATGATGTTGCAATGGTCAATTCACACCCCTAGTGGTCTCATAGCTGACGTGGCCAGA 291
Qy 91 GluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleValLeuAspAlaGln 110
Db 292 GAAATAATCAAAACGCATGATGTGTTTCTCCACCGCGCACAACTACAGCTGCTGCTAA 351 -
Qy 111 IleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAspHisTrpArgGln 130
Db 352 ATCTTTGGTTATGATGCAAAAGATGTGGCTTTCTGTGTTACTCCGCGAAGAGTGGAGACAA 411
Qy 131 MetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGlnSerSerArgLeu 150
Db 412 AAGATAAAGACATGTAAGGTTGAGCTTATGAGTCTGAAGAAGGTGCGGTGTTTTCATCC 471
Qy 151 IleArgGluGluGluMetGluAspAlaIleThrPheLeuArgSerLysAlaGlySerPro 170
Db 472 ATTAGACAAGAAGTTGTTACAGAGTTGGTCAAGCTATAGGTGAAGCGCTGTGTAGTGAA 531
Qy 171 -----ValAsnIleThrLysIleIleTyrGlyIleIleIleSerIleMetIleArg 187
Db 532 AGACCATGTGTGAATCTGACTGAGATGCTGATGCGAGCATCGAACGACATGTGTGTAGA 591
Qy 188 ThrSerValGly---AsnCyLysGln-----LysGluArgLeuLeu 200
Db 592 TGTGTTCTTGACCGGAAGTGTGATGATCGATGTCATGTCGTGGTGTGGCAGTAGCAGCTTGA 651
Qy 201 SerValAlaAspAlaValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPhePro 220
Db 652 CGGTTGGGAAGAAAGATTATGAGACTATTATCGGCTTTTCAGCGTGGGTGATTTCTTCCT 711
Qy 221 ThrTrpLysLeuLeuHisTyrIleIleGlyAlaGluSerLysProArgArgLeuHisGln 240
Db 712 TCGTTGGGTTGGGTTGACTATCTGACTGGCTTAATTCAGAGATGAAACACAGTTTCTC 771
Qy 241 GluIleAspAspIleLeuGluGluIleLeuAsnGluHisLysAlaAsnLysPheProGlu 260
Db 772 GCAGTAGATGTTCTTCCTTGATGAGTAATTCAGAACACAGAGCAGCAGTAACAAG---AAG 828
Qy 261 AlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValProValPro 280
Db 829 AATGATGACTTCTTGGGGATACTTCTTCAACTTCAAGAATGTGGGAGGCTTGACTTTCAG 888
Qy 281 ValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySerGluThr 300
Db 889 CTCGACCGAGATAACCTCAAGCAATCTCTAGTGGACATGATAATAGGTGGGAGTAGACACT 948
Qy 301 ThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGluLeuArg 320
Db 949 ACTTCAACAACCTCTAGTAATGGACTTTTCGGGAGTTCTTAGAAATCCAAATACCATGAAG 1008
Qy 321 LysAlaGlnGluGluValArgGlnValPheGlyGluMetGlyLys-----ValAspGlu 338
Db 1009 AAAGCTCAAGAGAGGTAAAGAGAGTGTGGGAATCAATTCCAAAGCAGTAGTCTGATGAA 1068
Qy 339 SerArgPheHisAspLeuLysPhePheLysValValLysGluThrLeuArgLeuHis 358
Db 1069 AATTGTGTAATCAATGAACCTCTTGAATGTGTAGTCAAGAANAACCTTTGAGATTACAT 1128
Qy 359 ProProVal---ValLeuIleProArgGluCyAsArgGluThrThrArgIleAspGlyTyr 377
Db 1129 CCACCCCTCTCTTTGATTGTCGAGACATCATCAAGTGTAAATAAGAGGGTAC 1188
Qy 378 GluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsn 397
Db 1189 GATATTCGCGCAAAACAATGTTATCAATGTCATGGCGCATCCAGAGGGATCCTGAA 1248
Qy 398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCyAlaIleAspTyr 417
Db 1249 TTATGGGATGATCCTGAAGAAATTTATCCGAAAGATTTGAAACTAGCCAAAGTTGATCT 1308
Qy 418 LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIle 437
Db 1309 AATGACAAGATTTTCAATTAATTCGTTCCGTTATGGGAGAGGGGATGCCCTGCAATG 1368
Qy 438 ThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuTyrHisPheAsnTrp 457
Db 1369 TCATTGACCTGGCTTCAACTGAGTATGTTCTGTAATCTCTTTGTTGTTGTTCAATTGG 1428
Qy 458 GluLeuAlaAspGly-----IleThrProGlnThrLeuAspMetThrGluAlaIleGly 475
```

Db 1429 AATATGTCGAGTCTGGACGTATATTGATGCACAAACATTGACATGAGTGAGACAAATGGA 1488
 QY 476 GlyAlaLeuArgLysLysIleAspLeuLysLeuIleProIleProTyrGln 492
 Db 1489 CTCACGTCTCAGTAAGAAGTACCACCTTCATCTTGAACCAACCATATAAA 1539

RESULT 7

US-10-953-349-14590
 ; Sequence 14590, Application US/10953349
 ; Publication No. US20060107345A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ALEXANDROV, Nickolai et al.
 ; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
 ; FILE OF INVENTION: ENCODED THERBY
 ; FILE REFERENCE: 2750-1579PUS2
 ; CURRENT APPLICATION NUMBER: US/10/953,349
 ; CURRENT FILING DATE: 2004-09-30
 ; NUMBER OF SEQ ID NOS: 40252
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 14590
 ; LENGTH: 1737
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 US-10-953-349-14590

Alignment Scores:

Pred. No.: 1-79e-79 Length: 1737
 Score: 828.50 Matches: 184
 Percent Similarity: 57.3% Conservative: 96
 Best Local Similarity: 37.6% Mismatches: 184
 Query Match: 32.2% Indels: 25
 DB: 6 Gaps: 9

US-10-759-813-2 (1-500) x US-10-953-349-14590 (1-1737)

QY 6 LeuSerPheProSerIleLeuIleSerPheLeuLeuValLeuIleLeuValValMet 25
 Db 170 CTAACCTCTTCCTTTGTTGTTGTTCTCTTC-----AAATACCGT 211
 QY 26 ArgLeuTrpLysGlnSerProProGlyProTrpLysPheProIleIleGlyAsn 45
 Db 212 AAAACCTTCAAGAAATCCACCCTTCCACCAGGTCTTAGAGGCTTCCCAATAATAGGGAAT 271
 QY 46 LeuProHisLeuLeuLeuTrpSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGln 65
 Db 272 CTT---CATCACTAGATAGCTCTCTCT---CATGACGATGCGATCTCTCNAAG 325
 QY 66 IleTyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValIleSerSer 85
 Db 326 AAATATGGTCCCTTATTTTCCCTTCAATTTGGATTAAAGCCAGCCATAGTTGTTCTCT 385
 QY 86 AlaGluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIle 105
 Db 386 CCAAAATGCGCCAAAGAGTAATGAAGACAAATGACCTTTGAGTGTGTGGACGACCTAAA 445
 QY 106 ValLeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGly 125
 Db 446 TTACTAGGCCAACAGAACTACCTATATATGGGTAGACATGGGATTTTCTCATACGAT 505
 QY 126 AspHisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysVal 145
 Db 506 AGTTATTGGAGAAATCAGAAAAATTTGTGTGTCCATGTCTTAGCTCCAAACGCTGTC 565
 QY 146 GlnSerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSer 165
 Db 566 CAARGCTTTACCTCAATAAGACATTTTGAGGTCAAGCAGATGATAAAAAATATCTAGG 625
 QY 166 LysAlaGlySerPro-----ValenIleThrLysIleIleTyrGlyIleIleSer 183
 Db 626 CATGCTCACTTCAAAAGTTACAAATTTAAATGAAGTCTTATATCCCTTTATAGTACT 685
 QY 184 IleMetIleArgThrSerValGlyAsnCysLysGlnLysGlu----- 197

Db 686 ATTGTATGTAAGATTCTTTTGGGAGAGAGGTATGTAAGAGGAGGATCTGAAGGTAGCAGA 745
 QY 198 -----ArgLeuLeuSerValAlaAspAlaValAsnGluAlaAlaThrSerPheGlyThr 215
 Db 746 TTCCATAAGTTGTTCATAGTGTGAAGCTATG-----TTGGGTAACTTCTTTGTT 796
 QY 216 AlaAspAlaPheProThrTrpLysLeuLeuHisTyrIleIleGlyAlaGluSerLysPro 235
 Db 797 TCAGATTATATTCCTTTTCATGGTTGGATTGATAAACTCAGGGACTGGATGACGCTCT 856
 QY 236 ArgArgLeuHisGlnGluIleAspIleLeuGluGluIleLeuLeuHisGluHis----- 253
 Db 857 GAACGCAATTTCAAGGAGATGATAGTTCTTCAAGAAGCCATTCATGATGAACACATGAAT 916
 QY 254 LysAlaAsnLysProPheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLys 273
 Db 917 TCTAAGAAAAAACTCCAGAGGAAGAGGATTTAGTTGATGTCTTACTTCAACTGAAGAG 976
 QY 274 AsnGlyAsnValProValProValThrAsnGluSerIleLysAlaSerValLeuGlnMet 293
 Db 977 AACACACGTTTCCCATAGATCTCACAATGATACATCAAGCAGTGTCTTGAACCTTA 1036
 QY 294 PheThrAlaGlySerGluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMet 313
 Db 1037 CTTGTAGGAGCTACAGGTACAACCTGAAGTCACAACAATCTGGGCCATGACTGAACATA 1096
 QY 314 LysAsnProThrGluLeuArgLysAlaGlnGluValArgGlnValPheGlyGluMet 333
 Db 1097 AAAAATCCAAAGCATTTATGAAGAAAGTGAAGAAGAAATTAGAGGCTTAAAGTGTGAAGAA 1156
 QY 334 GlyLysValAspGluSerArgPheHisAspLeuLysPhePheLysValValValLysGlu 353
 Db 1157 GATTTCCTTGATGAGATGATATTCAAAGTTTCTCTTAAGGCGAGTGAATAAAGAG 1216
 QY 354 ThrLeuArgLeuHis---ProProValValLeuIleProArgGluCysArgGluThrThr 372
 Db 1217 ACATTAGATTGCTACCTTACCAGCACCCTACTTATACCAAGAGAAACAAATAAAATGTC 1276
 QY 373 ArgIleAspGlyTyrGluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIle 392
 Db 1277 ATGTTAGATGGCTACGAAATTCAGCCAAAGACATTTATATGTGTAATCTTTGGGCAAT 1336
 QY 393 GlyArgAspProAsnThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAsp 412
 Db 1337 CATAGAGATCCTAAGCTTGGAAAGATCCAGAAAGATTATACCTGAGAGGTTTTTAAAT 1396
 QY 413 CysAlaIleAspTyrLysGlyThrPheGluLeuValProPheGlyAlaGlyLysArg 432
 Db 1397 TCGCAGCATAGATCTTTATGCAAGATTTTGAGTTTATTCATTCGTCGTGCTGCTAG 1456
 QY 433 IleCysProGlyIleThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeu 452
 Db 1457 TTGTGCTCTGATGAATATGTCATTTGCTGTCATTTGATGATCTTATATCTTCTTCT 1516
 QY 453 TyrHisPheAsnTrpGluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGlu 472
 Db 1517 TATTCCTTTGACTGGGAATGGCCCAAGGATGAAGAAAGAACAGACATTCAT---ACTGAA 1573
 QY 473 AlaIleGlyGlyAlaLeuArgLysLys 481
 Db 1574 GTGTTGCCAGGAGTTTACACAGCATGAAG 1600

RESULT 8

US-10-953-349-2239
 ; Sequence 2239, Application US/10953349
 ; Publication No. US20060107345A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ALEXANDROV, Nickolai et al.
 ; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
 ; FILE OF INVENTION: ENCODED THERBY
 ; FILE REFERENCE: 2750-1579PUS2
 ; CURRENT APPLICATION NUMBER: US/10/953,349

```
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2239
; LENGTH: 1617
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-2239

Alignment Scores:
Pred. No.:      2,43e-78      Length:      1617
Score:          817.50      Matches:      177
Percent Similarity: 55.9%      Conservatives: 111
Best Local Similarity: 34.4%      Mismatches: 192
Query Match:      31.8%      Indels:      35
DB:                6      Gaps:      10

US-10-759-813-2 (1-500) x US-10-953-349-2239 (1-1617)

QY 1 MetGluGlnLysAsnLeuSerPheProSerIleLeuIleSerPheLeuLeuValLeuIle 20
Db 10 TTAATAACAAATAACATAACAAATGATAGCATTTGGATTCTATCATCTTCTTCTTCATC 69
QY 21 LeuValValMetArgLeuTrpLysGlnAsn----- 32
Db 70 ATCTTTCTTCTTCGCGCCCTTCAACGCAAAATCATGGGAACATCGGAAATACCA 129
QY 33 ProProGlyProTrpLysPheProIleGlyAsnLeuProHisLeuLeuLeuThr 52
Db 130 TCTCCTCTCTGGT-----TTTCCATCATCGGAATTTATCATCAGCTC----- 171
QY 53 SerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIleTyrGlyProValMetSer 72
Db 172 GGAGAAATACAAATCAGCTCTCTATGGAGGCTCTCAAGAAGATATGGCCCTGTGTGCTT 231
QY 73 LeuGlnIleGlyGlnValSerAlaValIleSerSerAlaGluAlaAlaLysGluVal 92
Db 232 TTGAAGCTTGGAAAGATCCCCACATCATACATCTTCTCATCAGAACACAGAAACAGCT 291
QY 93 MetLysThrGlnAlaAspAlaPheAlaGlnArgProIleValLeuAspAlaGlnIleVal 112
Db 292 CTAAAGAGATATGATCTCCATTTGTTAGCCGCTCTCTTAGCAGGAGGAGAGAGCTC 351
QY 113 PheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAspHisTrpArgGlnMetLys 132
Db 352 TCTTACAACAATCTAGATATGCTTCATCTCCTTATAATGATATATGGAAGGAACATAAGG 411
QY 133 LysIleTrpIleLeuGluPheLeuSerAlaLysValGlnSerSerArgLeuIleArg 152
Db 412 AAGCTCTGAGTCAAGAACTCTTTAGTGCTATAATAAATCAATCATCATCAACCTATTAG 471
QY 153 GluGluGluMetGluAspAlaIleThrPheLeu-----ArgSerLysAlaGlySerPro 170
Db 472 GACGAGGAGGTCAAGAAATTTATCGATTCAATCGCGGAATCATCTTCTTAAAGAAATCCG 531
QY 171 ValAsnIleThrLysIleTyrGlyIleIleSerIleMetIleArgThrSerVal 190
Db 532 GTTAACCTTGAGCAAGACGTTCTTCTGTTAACTACAAGTGATGATGCAAGGCAGCATTT 591
QY 191 -----GlyAsnCysLysGlnLysGluArgLeuSerValAlaAspAla 205
Db 592 GGTGTGAGTTTGGAGGAAGTGTGCTCAACAGTGATAGATTCAATAAGTTAGTCGAGAT 651
QY 206 ValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPhePro-----ThrTrpLys 223
Db 652 ACTTTCGAGATGTTGGGAAGCTTCTCGCCTCAGATTTTATTCGCTATGTCGAGTGG--- 708
QY 224 LeuLeuHisTyrIleIleGlyAlaGluSerLysProArgArgLeuHisGlnGluLeuAsp 243
Db 709 ATCATCGATAAGTTCAATGTGTTTACAGGGTGGAGAAAGAAAGCTTTTCGAGATCTGAT 768
QY 244 AspIleLeuGluGluIleLeuAsnGluHisLysAlaAsnLysProPheGluAlaAspAsn 263
```

```
Db 769 CGGTTCTATGAACAAATTTTGTGATCTGCATGAAGAGAGAAAGTAGGAGTGAAGAT 828
QY 264 LeuMetAspValLeuLeuAsnLeuGlnLys-----AsnGlyAsnVal 277
Db 829 TTAGTGGATGTCTCTTGAGGTTGGAGAAAGAAATTTGTTGTTGGAATGGCAAG--- 885
QY 278 ProValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGly 297
Db 886 -----CTCACAAAGAAATCATATCAAGCAATTTTGTAGTGAACATCTTTTAGGAG 936
QY 298 SerGluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThr 317
Db 937 ATCGATATCTTCTGCAATAACATGACATGGCAATGGCAGAACTTGTCTAAAAACCTAGA 996
QY 318 GluLeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAsp 337
Db 997 GTGATGAAGAAAGTTCAAGCAGAAATCAGAAACCAAAATCAAGAACAAAGAAATCAGC 1056
QY 338 GluSerArgPheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeu 357
Db 1057 TTTGATGACATGATAAGCTGAGTACTTGAATAATGGTGATCAAGAAACATGGAGGTTA 1116
QY 358 HisProProVal---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGly 376
Db 1117 CATCTCCAACACCTCTCTCTCCCAAGAGATGTAATCACTGAATTTGAGATCAATGGC 1176
QY 377 TyrGluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspPro 396
Db 1177 TACAGATTCCTGCCAAACACACGCTTCATGTGAATGTTGGGCTATCGGCGGTATCCT 1236
QY 397 AsnThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAsp 416
Db 1237 GATACATGGAAGATCCAGAAATGTTCTCCCGAAAGGTTTAAATGATAGTAACATTGAT 1296
QY 417 TyrLysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGly 436
Db 1297 GCAAAAGGACAGAACTTTGAGTTGTTGTCGTTGGAGTGTGAGGAGAAATTTGCTCTGA 1356
QY 437 IleThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsn 456
Db 1357 CTGTACATGGGAACAACAATCGTGGAGTTTGGCCTAGCTAATATGTTGTATCATTTTGT 1416
QY 457 TrpGluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGly 476
Db 1417 TGAAGTTTACCAGAGGCTGTAGTCGAGATATCGATATGGAAGAAAGCTCTGAGACTT 1476
QY 477 AlaLeuArgLysLysIleAspLysLeuIleProIleProTyr 491
Db 1477 ACTGTGAGCAAAAAGTGAGCTTGTACTTGTTCAGTGAAGTAT 1521

RESULT 9
US-10-953-349-4847
; Sequence 4847, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4847
; LENGTH: 1586
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-4847

Alignment Scores:
Pred. No.:      3.02e-78      Length:      1586
Score:          816.50      Matches:      193
Percent Similarity: 57.4%
```



```
Qy 49 LeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIleTyrGly 68
Db 205 ATC-----GGGCCACATCGGCACCGGAGCTGGCTGGCTGGCCCGCGGCGATGG 255
Qy 69 ProValMetSerLeuGlnIleGlyGlnValSerAlaValIleSerSerAlaGluAla 88
Db 256 CCTGTGATGATGCTCGCGCTGGCATGGTGGCGACGGTGGTGTGACATCCGACAGAGCG 315
Qy 89 AlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleValLeuAsp 108
Db 316 GCGCGGGAAGCCCTCAAGACCCAGATGATGATGATGATGATGATGATGATGATGATGAT 375
Qy 109 AlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAspHisTrp 128
Db 376 CCAGGGCTGTGCTTACGGCTACAAGCAGCTGGCTTTCGCCGTGGAGGAGTACGTC 435
Qy 129 ArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGlnSerSer 148
Db 436 CGCGAGATGCCAAGTTGTTTCATCATCGAGTGTCTCAGCAGCGCGCGCTGCAGTCTGCC 495
Qy 149 ArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLysAlaGly 168
Db 496 TACTACGGGAGGATGCAGATGCACAGCTGATTTGAGACCTTACCGTGGTGGGCGG 555
Qy 169 SerProValAsnIleThrLysIleIleTyrGlyIleIleSerIleMetIleArgThr 188
Db 556 AACCGGTACCCCTGGAGGCCACATCTTCGCCACCATGGACCGGATTTGGGCTGTTC 615
Qy 189 SerValGlyAsnCys-----LysGlnLysGluArgLeuLeuSerValAlaAsp 204
Db 616 CGGTTCGGCGAGAGTACGCGGGGAGCAGTTCAAGGGGCGAGTTGTGCCCTTACTCAAC 675
Qy 205 AlaValAsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPhePro-----ThrTrp 222
Db 676 GCACCATGGACATGTTGGTGTGCTTCTGCCAAGACTTCTTCCCAAGCGCGCTGGC 735
Qy 223 LysLeuLeuHisTyrIleIleGlyAlaGluSerLysProArgArgLeuHisGlnGluIle 242
Db 736 CGCTCATCGCGCATCACCGGCTCAAGGCCACCGCCAGAGGGTCTTTTCGTGAGTTC 795
Qy 243 AspAspIleLeuGluGluIleLeuAsnGluHisLysAlaAsnLysProPhe-----Glu 260
Db 796 GACCGCTTCTTTGAGCATATCATCAGCAGTGCATGGGAGAAAGGCCACCGCGCGCGC 855
Qy 261 AlaAspAsnMetAspValLeuLeuAsnLeuGlnLysAsn-----GlyAsnValPro 278
Db 856 GGATCGGACCTGGTCAGAGCTGTGGACATCATGAAGAGCCCGCGCGCTCCGCTGCA 915
Qy 279 ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer 298
Db 916 GGAACCTTCACAGAGATCATGTCAAGGCCATTTCTCATGAACACATTCATTGGTAGCAT 975
Qy 299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
Db 976 GACACCAACACAGTACCATACTTGGCAATGGCAGAGCTGATCCGAAACCAAGGGTT 1035
Qy 319 LeuArgLysAlaGlnGluGluValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
Db 1036 CTGCAAGGGGCAACTTGATGATTTAGACTGCCGCTGGAGGAGGAGGAGTGCACCA 1095
Qy 339 SerArgPheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHis 358
Db 1096 GCTGACATGCCCCAAAATGAGTACCTGAGGATGGTGTGTCCGAGACCTACGGCTGCAT 1155
Qy 359 ProPro-----ValValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyr 377
Db 1156 CCCCAGCAACACTACTCTGATCCGAGAGAGAGCGCTCGCGCGCATCCAGGTGGCTGGCTAC 1215
Qy 378 GluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyValArgAspProAsn 397
Db 1216 GACATCCCGACCAAGACCCAGGTGCTATTGTTCATGTGGTGGCCATCAGCAGGAGACCTTCC 1275
```

```
Qy 398 ThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyr 417
Db 1276 GCGTGGGAAGGACCCGAGGAGTTTAAACCGGAGCGGTTCCAGGACACTGACGTGACTTC 1335
Qy 418 LysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIle 437
Db 1336 AACGGCAGCCACTTCGAGTTTCATCCCTTCGCGCGGGCGCGCGGATCTGCCCGGGCTG 1395
Qy 438 ThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrp 457
Db 1396 GCCATGGGGTGGCAACCGCGAGTACATCTTGCACACTTGTCTACTGCTTCAACTGG 1455
Qy 458 GluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyAla 477
Db 1456 CGCTGTCCTCAACGGGTGAGCGCGGAGGTGTGAACATGGAGGAGGAAGGGGTGCTCACT 1515
Qy 478 LeuArgLysLysIleAspLeuLysLeuIleProIleProTyr 491
Db 1516 TACCGGAAGAAGACCTCACTCATCTGCTGTGCAACACCTTAC 1557

RESULT 11
US-10-953-349-21349
; Sequence 21349, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 21349
; LENGTH: 1636
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-21349

Alignment Scores:
Pred. No.: 1,45e-76 Length: 1636
Score: 801.00 Matches: 171
Percent Similarity: 56.0% Conservative: 114
Best Local Similarity: 33.6% Mismatches: 184
Query Match: 31.1% Indels: 40
DB: 6 Gaps: 11

US-10-759-813-2 (1-500) x US-10-953-349-21349 (1-1636)

Qy 10 SerIleLeuIleSerPheLeuLeuValLeuIleLeuValValMetArgLeuTrpLys 29
Db 118 ACAATCTTAATCTCTGTCCTCCACTAACAATACTCTCTGCTGCTTATCTCGTACCGGT 177
Qy 30 LysGlnAsnPro---ProProGlyProTrpLysPheProIleIleGlyAsnLeuProHis 48
Db 178 CCAAGACCGCCTATTCACACGAGCGCTTCCCAATCATAGGAAC-----228
Qy 49 LeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIleTyrGly 68
Db 229 ATGTTTATGATGGACCACTAACCCACCGCGGTCTCGCCAACTCGCCAAACACTACGCG 288
Qy 69 ProValMetSerLeuGlnIleGlyGlnValSerAlaValIleSerSerAlaGluAla 88
Db 289 GGAATCTTCCACCTTCGATGGGTTCCTCCACATGTCGCGCATCTCCGACCCCGATGCC 348
Qy 89 AlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleValLeuAsp 108
Db 349 GCGGACAGGTTCCTCAAGTCCAGACAACTCTTTTCAACCGCCCGCCACCATCGCC 408
Qy 109 AlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAspHisTrp 128
Db 409 ATCAGCTACCTCACCTACGACCGCGCCGACATGGCTTCGCGCCACTACGCGCCCTTCTCG 468
```

```
QY 129 ArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGlnSerSer 148
Db 129 ArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGlnSerSer 148
QY 469 CGCCAGATGAGGAACTCTGCGTCATGAAGCTCTTCAGCGCGAAGCGCGCGAGTCTCTG 528
Db 469 CGCCAGATGAGGAACTCTGCGTCATGAAGCTCTTCAGCGCGAAGCGCGCGAGTCTCTG 528
QY 149 ArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLysAlaGly 168
Db 149 ArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLysAlaGly 168
QY 529 CAGTCGCTTCGC---GACGAGGTCGACTCGCGCGTCGCGCGCTCGCCACACGCGTCGGA 585
Db 529 CAGTCGCTTCGC---GACGAGGTCGACTCGCGCGTCGCGCGCTCGCCACACGCGTCGGA 585
QY 169 SerProValAsnIleThrLysIleIleTrpGlyIleIleSerIleMetIleArgThr 188
Db 169 SerProValAsnIleThrLysIleIleTrpGlyIleIleSerIleMetIleArgThr 188
QY 586 AAACCGCTCAATCCGGAATGATGTTAACTCACCAGAACATCATCTACCGCGCC 645
Db 586 AAACCGCTCAATCCGGAATGATGTTAACTCACCAGAACATCATCTACCGCGCC 645
QY 189 SerValGlyAsnCysLysGln-----LysGluArgLeuLeuSerValAlaAspAlaVal 206
Db 189 SerValGlyAsnCysLysGln-----LysGluArgLeuLeuSerValAlaAspAlaVal 206
QY 646 CGGTCGCGTTCGAGTTCGCGAGGAGTCAAGACGATTTTCATTAATAATATTCAGGAGTTC 705
Db 646 CGGTCGCGTTCGAGTTCGCGAGGAGTCAAGACGATTTTCATTAATAATATTCAGGAGTTC 705
QY 207 AsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeuHis 226
Db 207 AsnGluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeuHis 226
QY 706 TCTAAGCTCTTTGGCGCGTTAATATTGCGGATTTTATACCC----- 747
Db 706 TCTAAGCTCTTTGGCGCGTTAATATTGCGGATTTTATACCC----- 747
QY 227 TyrIle-----IleGlyAlaGluSerLysProArgArgLeuHisGlnGlu 241
Db 227 TyrIle-----IleGlyAlaGluSerLysProArgArgLeuHisGlnGlu 241
QY 748 TACCTCGCGCGGTAGATCCACAGGTTTGAATCCAGACTCGCTAGGCGACGTGCGCG 807
Db 748 TACCTCGCGCGGTAGATCCACAGGTTTGAATCCAGACTCGCTAGGCGACGTGCGCG 807
QY 242 IleAspAspIleLeuGluIleLeuAsnGluHisLysAlaAsnLys----- 257
Db 242 IleAspAspIleLeuGluIleLeuAsnGluHisLysAlaAsnLys----- 257
QY 808 CTCGATAGTTCATTGATAAGATCATCGACGACGTCGACAGAGAAGATAATATATCAG 867
Db 808 CTCGATAGTTCATTGATAAGATCATCGACGACGTCGACAGAGAAGATAATATATCAG 867
QY 258 -----ProPheGluAlaAspAsnLeuMetAspValLeu----- 268
Db 258 -----ProPheGluAlaAspAsnLeuMetAspValLeu----- 268
QY 868 AGCAGTGAAATTTGGTATGCGTGAACGAGCATGCTGGATGAGTGTGCGGTCTACGCG 927
Db 868 AGCAGTGAAATTTGGTATGCGTGAACGAGCATGCTGGATGAGTGTGCGGTCTACGCG 927
QY 269 -----LeuAsnLeuGlnLysAsnGlyAsnVal-----ProValProValThr 282
Db 269 -----LeuAsnLeuGlnLysAsnGlyAsnVal-----ProValProValThr 282
QY 928 GAGGAGGCGAAGTTGAATCGGACGATAATTTGCAGACCTCTATCAGACTCACT 987
Db 928 GAGGAGGCGAAGTTGAATCGGACGATAATTTGCAGACCTCTATCAGACTCACT 987
QY 283 AsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySerGluThrThrSer 302
Db 283 AsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySerGluThrThrSer 302
QY 988 AAGGATAAATCAAGGCTATCATATGAGCGTGTGTTTCGAGGCTCGGAACGTTGGCG 1047
Db 988 AAGGATAAATCAAGGCTATCATATGAGCGTGTGTTTCGAGGCTCGGAACGTTGGCG 1047
QY 303 LysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGluLeuArgLysAla 322
Db 303 LysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGluLeuArgLysAla 322
QY 1048 TCAGCAATCGAGTGGGTCTATCGGAGCTGATGAGAAGCCGACAGACCAAAAGCGGTC 1107
Db 1048 TCAGCAATCGAGTGGGTCTATCGGAGCTGATGAGAAGCCGACAGACCAAAAGCGGTC 1107
QY 323 GlnGluGluValArgGlnValPheGlyGluMetGlyLysValAspGluSerArgPheHis 342
Db 323 GlnGluGluValArgGlnValPheGlyGluMetGlyLysValAspGluSerArgPheHis 342
QY 1108 CAACAGAGCTCGCGATGTAGTGGCGCTGGACCGTGGCGGTGGAAGATCCGATTCGAG 1167
Db 1108 CAACAGAGCTCGCGATGTAGTGGCGCTGGACCGTGGCGGTGGAAGATCCGATTCGAG 1167
QY 343 AspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHisProValVal 362
Db 343 AspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHisProValVal 362
QY 1168 AAACCTCACTTATCTCAATGTGCCCTCAAGAGACCTTCGCGCTCCACCTCGATACCG 1227
Db 1168 AAACCTCACTTATCTCAATGTGCCCTCAAGAGACCTTCGCGCTCCACCTCGATACCG 1227
QY 363 LeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyrGluIleHisProAsn 382
Db 363 LeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyrGluIleHisProAsn 382
QY 1228 CTCCTCTCCAGACGCGCGGAGGACGCGCACCGTGGCGGTGCTACTTCGCTCCCGGAG 1287
Db 1228 CTCCTCTCCAGACGCGCGGAGGACGCGCACCGTGGCGGTGCTACTTCGCTCCCGGAG 1287
QY 383 ThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsnThrTrpSerGluPro 402
Db 383 ThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsnThrTrpSerGluPro 402
QY 1288 GCGCGTGTTATGATCAACCGTGGGCGCATTTGGAGGGAACAAGATTCCTGGGAGAACCC 1347
Db 1288 GCGCGTGTTATGATCAACCGTGGGCGCATTTGGAGGGAACAAGATTCCTGGGAGAACCC 1347
QY 403 GlyLysPheAsnProGluArgPhe---LysAspCysAlaIleAspTyrLysGlyThrThr 421
Db 403 GlyLysPheAsnProGluArgPhe---LysAspCysAlaIleAspTyrLysGlyThrThr 421
QY 1348 GAAACTTTTAAAGCCCGCGTTCCTTAAACCGCGCGTCCCGATTTCAAGGGAGGACAC 1407
Db 1348 GAAACTTTTAAAGCCCGCGTTCCTTAAACCGCGCGTCCCGATTTCAAGGGAGGACAC 1407
QY 422 PheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIleThrSerAlaIle 441
Db 422 PheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIleThrSerAlaIle 441
QY 1408 TTCGAGTTCATTCATTCGCGTGGGTGCGAGATCCGTCGCGGATGTTGCGGCTC 1467
Db 1408 TTCGAGTTCATTCATTCGCGTGGGTGCGAGATCCGTCGCGGATGTTGCGGCTC 1467
QY 442 ThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrpGluLeuAlaAsp 461
Db 442 ThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrpGluLeuAlaAsp 461
QY 1468 TAGCGGCTGGAATGCGGTGCTCATCTTCCTACTGCTTCACGTGGGAAATTCGAGAT 1527
Db 1468 TAGCGGCTGGAATGCGGTGCTCATCTTCCTACTGCTTCACGTGGGAAATTCGAGAT 1527
QY 462 GlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyAlaLeuArgLysLys 481
Db 462 GlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyAlaLeuArgLysLys 481
```

```
Db 1528 GGAATGAACCAAGTGAGATGGACATGGGTGACGTGTTCCGA-----CTCACCGCTCCA 1581
QY 482 IleAspLeuLysLeuIleProIlePro 490
Db 1582 AGGTCCACGCGACTCATTTGCTGTACCA 1608

RESULT 12
US-10-953-349-6466
; Sequence 6466, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6466
; LENGTH: 1637
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-6466

Alignment Scores:
Pred. No.: 2,37e-76 Length: 1637
Score: 799.00 Matches: 178
Percent Similarity: 56.3% Conservative: 106
Best Local Similarity: 35.3% Mismatches: 192
Query Match: 31.0% Indels: 28
DB: Gaps: 11

US-10-759-813-2 (1-500) x US-10-953-349-6466 (1-1637)

QY 13 IleSerPheLeuValLeuIleLeuValValMetArgLeuTrp----- 28
Db 26 ATCTCTTCTCTGTTGTTTCTCATCACCTTCGTTTCGTTTAATCTTTTGCCAGAAA 85
QY 29 -----LysLysGlnAsnProProGlyProTrpLysPheProIleIleGlyAsn 45
Db 86 ATCAACAGCATCAAAATGGAATCTTCTCCAAAGCCCTCCCAAGTTTCCGTCATCGGAA 145
QY 46 LeuProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGln 65
Db 146 TTACATCAGATT-----GGAAATTCCTCACAGTCTCAACATCTCGCCGAA 196
QY 66 IleTyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSer 85
Db 197 AGATACGACCGCTGTGATGCTTCTTACCTTTGGGTTTGTCCTTAACTGTGTCTCATCG 256
QY 86 AlaGluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIle 105
Db 257 AGAGAAGCGCGCTGAAGAAGTCTTAGAACTCATGACCTAGACTGTTGTCAGCGGCTAAG 316
QY 106 ValLeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGly 125
Db 317 CTGTGCGGACAAAGTTACTTCGCGGGAATTTAAAGATATCGGTTTACGCCATACGGT 376
QY 126 AspHisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysVal 145
Db 377 AACGAGTGAAGCGCGCGGTAAAGTTTGCCTTCGCTGAGCTTTTCTGTTGAAAAGGTT 436
QY 146 GlnSerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSer 165
Db 437 CAGTCCTTTAGGCATATCCGAGAGGAAGAAATCTTCTGCTCAAGCAACTGTCCGAA 496
QY 166 LysAla-----GlySerProValAsnIleThrLysIleIleTyrGlyIleIleSer 183
Db 497 TCTGCGGTGATGCTCTCCGCTCGATTTGAGCAAAATCCCTTTTCTGGCTAACCGCTAGT 556
QY 184 IleMetIleArgThrSerValGly-----AsnCysLysGlnLysGly 197
```



```
Db 557 ATCCCTTTTAGAGTTCCTTAGGACAGAAATTTTCACGAGACGATTTTATCGATAAAGAA 616
Qy 198 ArgLeuSerValAlaAspAlaValAaenGluAlaThrSerPheGlyThrAlaAasp 217
Db 617 AAGATCGAAGAGCTGCTTGAAGCTGAGACTGCCTAGCAGAGTTTACATTGTTCTGAT 676
Qy 218 AlaPheProThr-----TrpLysLeuLeuHisTyrIleIleGlyAlaGluSer 233
Db 677 TTCTTCCTGTTGCCGAGCTTGGATGG---CTCGTTGATTGGTTTTCGGGACAAACACAAG 733
Qy 234 LysProArgArgLeuHisGlnGluLeuAspIleLeuGluGluLeuLeuAsnGluHis 253
Db 734 AGACTCAACGATGTTTTTACAGCTCGATGCTCTGTTTCAACATGTCTATAGATGATCAT 793
Qy 254 ---LysAlaAsnLysProPheGluAlaAspAenLeuMetAspValLeuLeuAen---Leu 271
Db 794 TTAATCTCGNAGATCAAAAGACGACGACGACATCATCGATTCAATGTTGGATGTGATT 853
Qy 272 GlnLysAsnGly---AsnValProValProValThrAsnGluSerIleLysAlaSerVal 290
Db 854 CATAAACAAGGAGAGACAGTTCCTTAGAGCTCACAATAGATCATATCAAGGGGTTTCTC 913
Qy 291 LeuGlnMetPheThrAlaGlySerGluThrThrSerLysAlaThrGluTrpValMetAla 310
Db 914 GCGAATATATTTCTTGCAGGGATAGACACAGGGGCCATCACCATGATATGGCAGTGACG 973
Qy 311 GluLeuMetLysAsnProThrGluLeuArgLysAlaGlnGluGluValArgGlnValPhe 330
Db 974 GAGCTCGTTTAAACCCGAACTGATAAAGAAAGTTTCAAGGCGATATCGGAGAACACTT 1033
Qy 331 Gly---GluMetGlyLysValAspGluSerArgPheHisAspLeuLysPhePheLysLeu 349
Db 1034 GGCAGCAATAAGGAGAGAATACCAGGAAGATATCGAGAAAGTTCCTTACTTGAAGATG 1093
Qy 350 ValValLysGluThrLeuArgLeuHisProProValValLeuLeu---ProArgGluCys 368
Db 1094 GTAAATCAAGAAACAAATTCAGGTTTACACCCAGCAGCTCTCTTATACTTCCAAAGGGAACA 1153
Qy 369 ArgGluThrThrArgIleAspGlyTyrGluIleHisProAsnThrArgIleValValAaen 388
Db 1154 ATGGCTCACATCAAGTTCAAGGTATGATATTTCTCCCAAGAGGAGATCTTGTGTAAT 1213
Qy 389 AlaTrpAlaIleGlyArgAspProAsnThrTrpSerGluProGlyLysPheAsnProGlu 408
Db 1214 GTTTCGGCAATAGGAAGATCCCAAACTCTGCACAAACCCGAAAGAGTTTGACCTGAG 1273
Qy 409 ArgPheLysAspCysAlaIleAspTyrLysGlyThrThrPheGluLeuValProPheGly 428
Db 1274 AGGTTTATGGATAGCTTTGTTGATTATAGGGACAAACATTACGAGCTCTTACCATTTGG 1333
Qy 429 AlaGlyLysArgIleCysProGlyIleThrSerAlaIleThrAsnLeuGluTyrValIle 448
Db 1334 TCCGGTCGAAGATATGTCCTGGGATGCCAATGGGATTTGCTCGCTCGAATTTGGACTC 1393
Qy 449 IleAsnLeuLeuTyrHisPheAsnTrpGluLeuAlaAspGlyIleThrProGlnThrLeu 468
Db 1394 TTGAACCTTACTTCTTCGATTGGAAGTTGCCTGATGGATGACACATAAAGATATC 1453
Qy 469 AspMetThrGluAlaIleGlyAlaLeuArgLysIleAspLeuLysLeuIlePro 488
Db 1454 GATACTGAAGAGCTGGTACTCTTACAATAGTCAAGAAAGTACCTCTCAAGCTCGTTCCA 1513
Qy 489 IleProTyrGln 492
Db 1514 GTTCGAGTTTCAG 1525
```

RESULT 13

US-10-953-349-31583

; Sequence 31583, Application US/10953349

; Publication No. US20060107345A1

; GENERAL INFORMATION:

; APPLICANT: ALEXANDROV, Nikolai et al.

```
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THEREY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953.349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 31583
; LENGTH: 1675
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-953-349-31583
```

```
Alignment Scores:
Pred. No.: 8.08e-74 Length: 1675
Score: 775.50 Matches: 185
Percent Similarity: 55.1% Conservative: 99
Best Local Similarity: 35.9% Mismatches: 181
Query Match: 30.1% Indels: 51
DB: 6 Gaps: 12
```

US-10-759-813-2 (1-500) x US-10-953-349-31583 (1-1675)

```
Qy 13 IleSerPheLeuLeuValLeuValValMetArgLeuTrpLysLysGlnAen 32
Db 108 CTTCACATCGTCCTCTCTCTCCCTTCGCCGTCTCTCTGTGTATATCTCTCGTAGG 167
Qy 33 Pro-----ProProGlyProTrp 38
Db 168 CCCGCCCTCTGCGCAGCGCAGCATGGAGGACGAGGAGCATCTTCCCGCTCACCGCGC 227
Qy 39 LysPheProIleIleGlyAsnLeuProHisLeuLeuLeuThrSerAspLeuGlyHisGlu 58
Db 228 GGCCTCCCGCTGCTCGGC-----CACCTACATCTCTCGGCTCGTTCGCGCACCGG 278
Qy 59 ArgPheArgAlaLeuAlaGlnIleTyrGlyProValMetSerLeuGlnIleGlyGlnVal 78
Db 279 GCCCTGCGGTCTCTGGCCGCGCACACGCGCGGTCTCTGCTCTCGGCTCGCGCGCTG 338
Qy 79 SerAlaValValIleSerSerAlaGluAlaLysGluValMetLysThrGlnAlaAep 98
Db 339 CCGCGCGTGTGTCTCTCGCGCGCGCGCGAGAGAGGTATGAGGCGCCGCGCACCTG 398
Qy 99 AlaPheAlaGlnArgProIleValLeuAspAlaGlnIleValPheTyrAsnArgLysAep 118
Db 399 GCCTTCGCGAGCGCGCGCGCGCGAGCCCATGGCCGACCCGCTCTCTACGGCGC---GAC 455
Qy 119 ValLeuPheAlaSerTyrGlyAspHisTrpArgGlnMetLysLysIleTrpIleLeuGlu 138
Db 456 GTAGCGTTTCGCGCCCTAGCGCGAGTACTGGCGCCAGCGCGCGCGCGCTGTGTGTCCAC 515
Qy 139 PheLeuSerAlaLysLysValGlnSerSerArgLeuIleArg-GluGluGluMetGluAs 158
Db 516 CTCCTTCAGCGCGCTCCGATCTCTCTCTCGCGCGCTCCGCGGAGGAGGAGCGCCGCG 575
Qy 158 pAlaIleThrPheLeuArgSerLysAla-----GlySerProValAsnIleThrLysI 176
Db 576 GCTGTGTCGAGCGCGCTCCGCGAGCGCGCGCGCGCGCGCTGTGGACCTGTGTGAGCT 635
Qy 176 eileTyrGlyIleIleIleSerIleMetIleArgThrSerValGlyAsnCysLys---- 194
Db 636 CCTCTCGCTAC-GCCAACACGCTGTCTCGCGCGCGCGCGCTTCGGGACGACAGCGCGG 694
Qy 195 -----GlnLysGluArgLeuLeuSerValAlaAspAlaValAaenGlu 208
Db 695 CGGCTGTATGAGGAAGGCAACAGGAACGC-----GAGCTGAGGAAGGTGTCAACGA 748
Qy 208 uAlaAlaThrSerPheGlyThrAla-----AspAlaPheProThrTrpLysLeuLe 225
Db 749 CTTCCAGGAGCTGCTCGGCACCGCGCCTCTGGGGAGCTCTCGGCCCTGCTGGGTGGT 808
Qy 225 uHisTyrIleIleGlyAlaGluSerLysProArgArgLeuHisGlnGluIleAspAsp 245
Db 225 uHisTyrIleIleGlyAlaGluSerLysProArgArgLeuHisGlnGluIleAspAsp 245
```

```
Db 809 GGACGGCGTGAGGGGATGCGAGGGGAGATCAGACGGACGTTCAAGCGCGCTCGACGGTGT 868
QY 245 eLeuGluGluLeuAAsnGluHisLysAlaAAsnLysProPhe----- 259
Db 869 GCTCGAAGAGGTATCGGCGACCAACCCCGCGGCTCAAGCGCGCCAGACGAGACGGGAGA 928
QY 260 -----GluAlaAspAAsnLeuMetAAspValLeuLeuAAsnLeu---GlnLysAAsnG1 275
Db 929 TGACGGCGCGCATCACAGGATTCGTGGACGTGTGTGGACGTGAGCGATCTGATGA 988
QY 275 yAAsnValProValThrAAsnGluSerLysAlaSerValLeuGlnMetPheThr 295
Db 989 CGAAGCTGCGATCGCTCAGCAGCAGCCGAAATCAAGGCCCATATTTTGGACATGTTGCG 1048
QY 295 xAlaGlySerGluThrSerLysAlaThrGluTTPValMetAlaGluLeuMetLysAs 315
Db 1049 GGCGGCGCAGGACACGACGACGCGAATGAGATGGCCATGGCGAGGTCAATCAGCA 1108
QY 315 nProThrGluLeuArgLysAlaGlnGluGluValArgGlnValPheGlyGluMetGlyLy 335
Db 1109 CCACAGACAGATCGTAAGTCCAGGACGAGCTCAGGCGCGCGTCTCGGTGGCTCCGGCA 1168
QY 335 sVal---AAspGluSerArgPheHisAAspLeuLysPheLysLeuValValLysGluTh 354
Db 1169 CGTCATCACCGAGGACCATCGACAAAGCTGCACTACCTCAAGGCGGTGTCTAAGGAGAC 1228
QY 354 xLeuArgLysProProVal---ValLeuLeuProArgGluCysArgGluThrThrAr 373
Db 1229 GCTACGCGCTGCACCCACCATCCCGCTCTCTGTCGCCCCGAGAGCCCAAGACGACGCCGA 1288
QY 373 gLeAAspGlyTyRgLuHisProAAsnThrArgLysValValAAsnAlaTTPAlaIleG1 393
Db 1289 GATACTCGGCCACACGCTCCCGCGCGGCGACGCGGGTGGTCAATCAACGCGTGGGCGCTCG 1348
QY 393 yArgAAspProAAsnThrTTPSerGluProGlyLysPheAAsnProGluArgPheLysAAspCy 413
Db 1349 CCGGGACCGCGCGCGCTGGGAGCGCGCGAGGAGTTCGTACCGGAGAGGTTTCTCGACGG 1408
QY 413 sAlaIleAAspTyRgLyGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArg1 433
Db 1409 CGCGGTGGACTACAGGCGGAGGACTTCAGCTGATACCGTTCCGTGCTGGCGCGCGGG 1468
QY 433 eCysProGlyIleThrSerAlaIleThrAAsnGluTyRgValIleAAsnLeuLeuTy 453
Db 1469 GTGCCCCGGGTCTGGATTCGCGCGCGCGCGGTTGAGATGGCGCTGCGGAGTTGATGA 1528
QY 453 rHisPheAAsnTTPGluLeuAlaAAspGlyIleThrProGlnThrLeuAAspMetThrGluAl 473
Db 1529 CCATTTGCACTGGGAGCGCGCGGGCG-----TCGCTGGACATGCGCGAGGT 1576
QY 473 alLeGlyGlyAlaLeuArgLysLysIleAAspLeuLysLeuIle 487
Db 1577 GAACGGGCTCGCGCGCATCTCAAGTCCGCGCTCGCGCTGTC 1619

RESULT 14
US-10-759-813-2
; Sequence 12597, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 12597
; LENGTH: 1852
; TYPE: DNA
; ORGANISM: Glycine max
US-10-759-813-2
```

```
Alignment Scores:
Pred. No.: 3,78e-72 Length: 1852
Score: 760.50 Matches: 170
Percent Similarity: 54.5% Conservative: 100
Best Local Similarity: 34.3% Mismatches: 196
Query Match: 29.5% Indels: 29
DB: 6 Gaps: 10

US-10-759-813-2 (1-500) x US-10-953-349-12597 (1-1852)

QY 18 ValLeuLeuValValMetArgLeuTrpLysGlnAAsnPro---ProProGly 36
Db 119 ATCCTCATATATCGTATAATAAGTTTATAACACGCCCATCCCTCCCTCCACCGGGC 178
QY 37 ProTrpLysPheProIleLeuGlyAAsnLeuProHisLeuLeuThrSerAAspLeuGly 56
Db 179 CTTAAACCATATGGCCCATATAGTGGGAAACTTGCCACACATG-----GGCCGGGTGCCG 229
QY 57 HisGluArgPheArgAlaLeuAlaGlnIleTyRgProValMetSerLeuGlnIleGly 76
Db 230 CACCACCTCCCTCGCGGGTGGCCCGCATCCACGGCCCGTTGATGACACCTCCGCTCGGC 289
QY 77 GlnValSerAlaValIleSerSerAlaGluAlaLysGluValMetLysThrGln 96
Db 290 TTCGTCGATGTCGTGTTGCGGCTCGGCTCGGCGGACAGTCTTGAAGATTTCAT 349
QY 97 AlaAAspAlaPheAlaGlnArgProIleValLeuAAspAlaGlnIleValPheTyRgAAsnArg 116
Db 350 GACTCTTAATTTGAGCAGCGCGGCCACCGAACCGCGCGCGGAAATATATAGCTTTATACTAC 409
QY 117 LysAAspValLeuPheAlaSerTyRgLyAAspHisTTPArgGlnMetLysLysIleTTPile 136
Db 410 CAAGATCTGGTGTGGCCCTACGCGCCACGCTGGCGATTGCTCCGGAACATCACCTCC 469
QY 137 LeuGluPheLeuSerAlaLysLysValGlnSerSerArgLeuLeuArgGluGluMet 156
Db 470 GTTCACCTCTCTCCGCAAGCCCATGAATGAATTTAGACACTTGGCTCGGAGAGGTA 529
QY 157 GluAAspAlaIleThrPheLeuArgSerLysAlaGlySerProValAAsnIleThrLysIle 176
Db 530 GCTAGATTGACATGCACTTGGCAAGTTTCAGACACAAAGAGCTAGTGAATTTGGACAAATTA 589
QY 177 IleTyRgLyIleIleIleSerIleMetIleArgThrSerVal----- 190
Db 590 TTGAATGATGACACCACCAACGATTTAGCGAGGGCAATGATTCGACGACAGAGTGTAAAC 649
QY 191 -----GlyAAsnCysLysGlnLys---GluArgLeuLeuSerValAlaAAspAlaVal 206
Db 650 GATGCAATGGAGGGTGTGATCCTAGGGCTGATGAGTTTAAAGCTATGGTGTATGGAGGTG 709
QY 207 AAsnGluAlaAlaThrSerPheGlyThrAlaAAspAlaPheProThrTrpLysLeuLeuHis 226
Db 710 ATGGGTGTTGGCTGGAGTTTTCACATTTGTTGACCTTTATTCCTTCCTGGAGTGGCTAGAC 769
QY 227 TyRgIleLeuGlyAlaGluSerLysProArgArgLeuHisGlnGluIleAAspAlaLeu 246
Db 770 ---CTTCAAGGGGTGCAAGCTAAGATAAGAAAATTAACAAAGAGATTGATGATTTTAA 826
QY 247 GluGluLeuLeuAAsnGluHis-----LysAlaAAsnLysProPheGluAlaAAspAAsnLeu 264
Db 827 ACCAGCATTTATTTAGGAGGACAAATAACAGTTCTTCCAGAAGATGAGAAATCATAAGAATTC 886
QY 265 MetAAspValLeuAAsnLeuGln-----LysAAsnGlyAAsnValProValPro 280
Db 887 TTGAGTATCTTATTTGCTAAAGATGTGAGGATGATCATGTGAAGAAC-----CAT 937
QY 281 ValThrAAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySerGluThr 300
Db 938 CTCACCTGATACCGAGATTAAAGCACTACTCTTGAACATGTTTACCGCAGCACTGACACA 997
QY 301 ThrSerLysAlaThrGluTTPValMetAlaGluLeuMetLysAAsnProThrGluLeuArg 320
Db 998 TCATCAAGCACCCACAGATGGGCCCACTAATGCGGAATAATAAAAAAACCCCAAAATTTTGGCC 1057
```

```
QY 321 LysAlaGlnGluGluValArgGlnValPheGlyGluMetGlyLysValAspGluSerArg 340
DB 1058 AAATCCCAACAAGAGTGTGACACACCTGTGGTTCGAGATAGAGTGTCAAGAAGAGGAC 1117
QY 341 PheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHisProPro 360
DB 1118 TTGGCCCACTCCCATCTTCAACAGCTGTATCAAGSAAACCTTCAGACTCCACCCATCA 1177
QY 361 ValValLeu---IleProArgGluCysArgGluThrThrArgGileAspGlyTyrgluile 379
DB 1178 ACCCTCTTTCTGTGCCAGCTGCTCTGCCGAGAGTGTGTGAGATATTGGCTACCATC 1237
QY 380 HisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsnThrTrp 399
DB 1238 CCCAAGGTGTACACTCTTGTGTGAACATTGGGCCATAGCCCGAGACCCAAAAGAGTGG 1297
QY 400 SerGluProGlyLysPheAsnProGluArgPhe-----LysAspCysAlaIle 415
DB 1298 AATGACCCATTGGAGTTTAGGCCCGAAGGTTTCTTAGGTGGTGGAAGGCCGATGTT 1357
QY 416 AspTyrlsGlyThr-Thr-PheGluLeuValProPheGlyAlaGlyLysArgIleCysPro 435
DB 1358 GATGTAGGGCAATGACTTTGAGGTGATACCTTTTGGTGTGGACGAGAAATTGTGCT 1417
QY 436 GlyIleThrSerAlaIleThrAsnLeuGluTyrlValIleIleAsnLeuLeuTyrlHisPhe 455
DB 1418 GGACTTAGTCTTGGGCTTCAATGTTCAACTTCTCACTGCGAGCCTTGCCCATTCATT 1477
QY 456 AsnTrpGluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGly 475
DB 1478 GATTCGGAACTAGAAGATTGCATGAACCCAGAAAAGTTGAACATGGATGAAGCGTATGG 1537
QY 476 GlyAlaLeuArgLysLysIleAspLeuLysLeuIleProIlePro 490
DB 1538 CTAACCTTACAAGCGCGTGGCTTTTCAGTGCACCTTAGGCC 1582
```

RESULT 15

```
US-10-953-349-21323
; Sequence 21323, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 21323
; LENGTH: 1660
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-21323
```

```
Alignment Scores:
Pred. No.: 3,65e-72 Length: 1660
Score: 760.00 Matches: 181
Percent Similarity: 54.8% Conservative: 103
Best Local Similarity: 34.9% Mismatches: 184
Query Match: 29.5% Indels: 51
DB: 6 Gaps: 12
```

US-10-759-813-2 (1-500) x US-10-953-349-21323 (1-1660)

```
QY 11 IleLeuIleSerPheLeuLeuValValLeuIleLeuValVal-----ValMetArgLeuTrp 28
DB 16 GTTCTCTTCTATTATGTCTATGCTATGATCTCCATCTCCATTTGTTCTTTCATACCCACACAC 75
QY 29 LysLysGlnAsnPro-----ProProGlyProTrpLysPheProIleIleGlyAsn 45
DB 76 AAAGAACAACACTCCAAGAGGACCAACCGAGTCTCCACCTCTTCTCTCATCGGCAC 135
```

```
QY 46 Leu-----ProHisLeuLeuLeuThrSerAspLeuGlyHisGlu 58
DB 136 CTTCAACAACTCCACAACATCATCCCCACATCTCTGCTTATGGCAA----- 180
QY 59 ArgPheArgAlaLeuAlaGlnIleTyrglyProValMetSerLeuGlnIleGlyGlnVal 78
DB 181 -----CTGCGCAAACTCCAGGTCTCTCATGTCTGTTTCGGCTTCGGCCGCTG 228
QY 79 SerAlaValValIleSerSerAlaGluAlaLysGluValMetLysThrGlnAlaAsp 98
DB 229 CAACCGTCGTGGTTCATCGCCAGAAATCCCGAACAATCTTGAAACCCACGACCTC 288
QY 99 AlaPheAlaGlnArgProIleValLeuAspAlaGlnIleValPheTyrlAsnArgLysAsp 118
DB 289 AACTTCGCTTCAGCGCTCTCTGTCGGGCCGAGAAAGCTCTCTTACGACGGGTGGAC 348
QY 119 ValLeuPheAlaSerTyrglyAspHisTrpArgGlnMetLysValIleTrpIleLeuGlu 138
DB 349 ATGGCTTCGCAACCGTACGGCCCGTACGAGAGAAATGAAGAAACTCTGCATCGTTTCA 408
QY 139 PheLeuSerAlaLysLysValGlnSerSerArgLeuIleArgGluGluGluMetGluAsp 158
DB 409 CTCCTCAGCGCGCAACGCTTCGTCCTTCGCAATTCGAGAGACGAGGTTCAGAAA 468
QY 159 AlaIleThrPheLeu-----ArgSerLysAlaGlySerProValAsnIleThrLysIle 176
DB 469 ATGGTTCGAAACCTTCGAAACACGAACTTCGGGTACTGTCGTGAACCTTGACCGAACT 528
QY 177 IleTyrglyIleIleIleSerIleMetIleArgThrSerValGly-----AsnCys 193
DB 529 TTGATGCTTTCACGAACCTCTTTGATATGCAAGATCGCGTTGGGGAAGAAAGTTACCGTGT 588
QY 194 LysGlnLysGlu-----ArgLeuLeuSer 201
DB 589 GAGTACGAGGAAGTAGTTGTTGATGAGGTACTGGAACCCGAGGAGCAGGTTCGAGGT 648
QY 202 ValAlaAspAlaValAsnGluAlaIleThrSerPheGlyThrAlaAspAlaPhe---Pro 220
DB 649 CTGCTCAACGAGGCTCAAGCGTTGCTTCGAGTTTCTTTTCGGATTATTTCCGCT 708
QY 221 ThrTrpLysLeuLeuHisTyrlIleGlyAlaGluSerLysProArgArgLeuHisGln 240
DB 709 ATAGAAAGTGGTGTGATAGTAGTCGGAATCTATCGCGCTTGATAAAACGTTCAAG 768
QY 241 GluIleAspAspIleLeuGluGluLeuAsnGluHis----- 253
DB 769 GAGTTGGACGCTGTCTACGAACGATTCTATGATCATCATGATTCATGATTCGCAAGAGTGT 828
QY 254 ---LysAlaAsnLysProPheGluAlaAspAsnLeuMetAspValIleLeuAsnLeuGln 272
DB 829 AAAAAAGATAATGACAAACAAGAGTCAAGATATTATTTGATATTCTTCTCCAGCTACTT 888
QY 273 LysAsnGlyAsnValProValProValThrAsnGluSerIleLysAlaSerValLeuGln 292
DB 889 GATGATCTCTCTCACCCTTTGATCTCACTCTCGACCATATAAAGCCGTGCTCATGAAC 948
QY 293 MetPheThrAlaGlySerGluThrThrSerLysAlaThrGluTrpValMetAlaGluLeu 312
DB 949 ATCTTTATAGCAGGAACAGACCCGAGTTCGCGACATAGTTTGGGCAATGAATGCACCTG 1008
QY 313 MetLysAsnProThrGluLeuArgLysAlaGlnGluValArgGlnValPheGlyGlu 332
DB 1009 TTGAAGAATCCCAATGTGATGAGCAAGTTCAGGAGAGTGAAGAAATCTATTTCGGTGAC 1068
QY 333 MetGlyLysValAspGluSerArgPheHisAspLeuLysPhePheLysLeuValLys 352
DB 1069 AAAGATTTCATAAGCAAGATGATCTCGAAGCCCTTCCTTATCTCAAGCAGTGTGTAAG 1128
QY 353 GluThrLeuArgLeuHisProPro---ValValLeuIleProArgGluCysArgGluThr 371
DB 1129 GAGCATTAAGATTATTCACCCTTCACCACTACTCTTTTGGCAAGGGTAAACAATGGAACA 1188
```


This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 31, 2006, 06:41:26 ; Search time 283 Seconds
(without alignments)
4958.769 Million cell updates/sec

Title: US-10-759-813-2

Perfect score: 2574

Sequence: 1 MEQKNLSFSPISILIFLLVLI.....KIDKLPIPYQVSLGNSIS 500

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delopt 6.0 , Delext 7.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abs/ABSWEB_spool/US10759813/runat_31052006_060309_2699/app_query.fasta_1
-DB=issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=Bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abs02h -USER=US10759813 @CGN 1 1 204 @runat_31052006_060309_2699
-NCPU=6 -ICPU=3 -NO_MMAP -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:
1: /EMC_Celerra_SIDS3/ptodata/2/ina/1 COMB.seq:
2: /EMC_Celerra_SIDS3/ptodata/2/ina/5 COMB.seq:
3: /EMC_Celerra_SIDS3/ptodata/2/ina/6A COMB.seq:
4: /EMC_Celerra_SIDS3/ptodata/2/ina/6B COMB.seq:
5: /EMC_Celerra_SIDS3/ptodata/2/ina/7 COMB.seq:
6: /EMC_Celerra_SIDS3/ptodata/2/ina/H COMB.seq:
7: /EMC_Celerra_SIDS3/ptodata/2/ina/PCTUS COMB.seq:
8: /EMC_Celerra_SIDS3/ptodata/2/ina/PP COMB.seq:
9: /EMC_Celerra_SIDS3/ptodata/2/ina/RE COMB.seq:
10: /EMC_Celerra_SIDS3/ptodata/2/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1127.5	43.8	1781	3	US-09-499-302A-1
2	1119	43.5	1665	2	US-08-313-075A-29
3	1088	42.3	1666	2	US-07-912-900-24
4	1088	42.3	1666	2	US-08-285-309-24
5	1088	42.3	1666	2	US-08-502-046-24
6	993.5	38.6	1665	3	US-09-292-768-67
7	993.5	38.6	1665	3	US-09-292-768-69
8	992.5	38.6	1515	3	US-09-292-768-5

9	992.5	38.6	1665	3	US-08-881-784-8	Sequence 8, Appli
10	992.5	38.6	1665	3	US-09-292-768-3	Sequence 3, Appli
11	983.5	38.2	1762	3	US-08-881-784-5	Sequence 5, Appli
12	983.5	38.2	1762	3	US-09-292-768-1	Sequence 1, Appli
13	979.5	38.1	1762	3	US-09-292-768-65	Sequence 65, Appli
14	978.5	38.0	1762	3	US-09-292-768-63	Sequence 63, Appli
15	977.5	38.0	1691	3	US-08-948-564-3	Sequence 3, Appli
16	955.5	37.1	1762	3	US-09-172-339-5	Sequence 5, Appli
17	856.5	33.3	1929	3	US-09-380-420C-1	Sequence 1, Appli
18	856.5	33.3	1929	3	US-09-380-420A-1	Sequence 1, Appli
19	840	32.6	1838	3	US-08-948-564-1	Sequence 1, Appli
20	805.5	31.3	2013	3	US-09-615-192A-404	Sequence 404, App
21	801	31.1	1539	5	US-10-142-231-53	Sequence 53, Appli
22	801	31.1	1539	5	US-10-884-115-53	Sequence 53, Appli
23	776	30.1	1866	3	US-09-615-192A-103	Sequence 103, App
24	776	30.1	1866	3	US-09-169-789-103	Sequence 103, App
25	774.5	30.1	1883	3	US-08-991-677-3	Sequence 3, Appli
26	774	30.1	1657	3	US-08-948-564-11	Sequence 11, Appli
27	773	30.0	1737	3	US-09-142-108C-3	Sequence 3, Appli
28	767.5	29.8	1789	3	US-09-142-108C-1	Sequence 1, Appli
29	767	29.8	1824	3	US-09-142-108C-20	Sequence 20, Appli
30	755.5	29.4	1764	3	US-09-947-027-3	Sequence 3, Appli
31	747	29.0	1660	3	US-09-142-108C-16	Sequence 16, Appli
32	747	29.0	1838	2	US-09-091-432-1	Sequence 1, Appli
33	747	29.0	1838	3	US-09-387-663-1	Sequence 1, Appli
34	747	29.0	1838	3	US-09-214-139B-2	Sequence 2, Appli
35	744.5	28.9	1880	3	US-09-564-808-1	Sequence 1, Appli
36	743.5	28.9	1711	3	US-09-142-108C-5	Sequence 5, Appli
37	739.5	28.7	1748	3	US-09-142-108C-14	Sequence 14, Appli
38	739.5	28.7	1884	3	US-09-564-808-3	Sequence 3, Appli
39	733	28.5	1545	5	US-10-142-231-55	Sequence 55, Appli
40	733	28.5	1545	5	US-10-884-115-55	Sequence 55, Appli
41	731.5	28.4	1667	3	US-09-142-108C-22	Sequence 22, Appli
42	731.5	28.4	1835	3	US-09-564-808-5	Sequence 5, Appli
43	728.5	28.3	1730	3	US-09-672-785-3	Sequence 3, Appli
44	727	28.2	1663	5	US-10-146-374-1	Sequence 1, Appli
45	719.5	28.0	1737	3	US-09-126-420A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-499-302A-1
; Sequence 1, Application US/09499302A
; Patent No. 6369212
; GENERAL INFORMATION:
; APPLICANT: BOUNG-JUN, OH
; APPLICANT: MOON, KYUNG KO
; APPLICANT: YOUNG, SOON KIM
; TITLE OF INVENTION: A CYTOCHROME P450 GENE HIGHLY EXPRESSED IN THE
; FILE REFERENCE: 10324/P6443USO
; CURRENT FILING DATE: 2000-02-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1781
; TYPE: DNA
; ORGANISM: Capsicum annuum
US-09-499-302A-1

Alignment Scores:
Pred. No.: 1.15e-114 Length: 1781
Score: 1127.50 Matches: 214
Percent Similarity: 67.3% Conservative: 118
Best Local Similarity: 43.4% Mismatches: 148
Query Match: 43.8% Indels: 13
DB: 3 Gaps: 5

US-10-759-813-2 (1-500) x US-09-499-302A-1 (1-1781)

Qy 12 LeuileSerPheLeuValLeuValMetArgLeuTrpLys 30

Db 25 TTAGTTGCGTTTGTCTCTCTCCAGCATCTTCTCTCCAGCATCTTCTCTCAAAAAATGGAAAAACC 84
Qy 31 Gln-----AsnProProGlyProThrLysPheProIleIleGlyValAsnLeuProHis 48
Db 85 CAAAGCTAACTTACCTCTGTCGTCGTAATACCTTTTATTTGGAAGCTTACATCAC 144
Qy 49 LeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIleTyrGly 68
Db 145 TTGGCAGTGGCAGGTCCTCTCTCATCATGCGCTAAAAAATCTAGCAAAACTTTATGGG 204
Qy 69 ProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSerAlaGluAla 88
Db 205 CCGCTCATGCATACGATCGGGGAAATTCCTACCGTCATCATTTTCGTCGCCGGAATG 264
Qy 89 AlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleValLeuAsp 108
Db 265 GCGAGGAGTACTAAAACTCAGACCTCGCTTCGCAACGAGGCCGAACTTGTGGTG 324
Qy 109 AlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAspHisTrp 128
Db 325 GCTGACATCGCTCATATGATAGTACGGATATAGCATTTTCTCCATATGCTGAATCTGG 384
Qy 129 ArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValGlnSerSer 148
Db 385 AGGCAGATTTCGTAATAATTTGCATATCGAACTCCCTTAGTGCACCAAGATGTCAAATCTTT 444
Qy 149 ArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSerLysAlaGly 168
Db 445 AGCTCAATTCGCAGGATGAGCTGTCGATGATGGTCTCATCATACGAACCATGCCAAT 504
Qy 169 SerProValAsnIleThrLysIleIleTyrGlyIleIleSerIleMetIleArgThr 188
Db 505 TTTCCGCTCAACCTTACAGCAAAATATTTTGGTTTACAAGTTCCGTAACCTTGTAGATCA 564
Qy 189 SerValGlyAsn---CysLysGlnLysGluArgLeuLeuSerValAlaAspAlaValAsn 207
Db 565 GCTCTGGGAAATATGTGTCGACCAAGCAAACTGATATTTTCATGAGGGAATAATA 624
Qy 208 GluAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeuHisTyr 227
Db 625 TCATTGACAGGTGGATTAGTATTGCTGATTTTTCCTTACATGGAATAATGCTACATGAT 684
Qy 228 IleIleGlyAlaGluSerLysProArgArgLeuHisGlnGlnIleAspAspIleLeuGlu 247
Db 685 GTTGGTGGTTCAAAAACTAGACTGCTGAAGGCTCATCGTAAAAATCGATGAGATTTGGAA 744
Qy 248 GluIleLeuAsnGluHisLysAlaAsnLys-----ProPhe 259
Db 745 CATGTAGTAGAATGAGCACAACAGAAATCGAGCGGATGGCCAAAGGTAATGCGCAATTT 804
Qy 260 GluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValProVal 279
Db 805 GCGGTTGAAGATTGATCGATGTTTTGCTAAGGGTTCGAGAAAGTCGAGAAGTTCAAAT 864
Qy 280 ProValThrAsnGluSerLysAlaSerValLeuGlnMetPheThrAlaGlySerGlu 299
Db 865 TCCATCAGCGATGACAATATCAATCAATATTAGTCGACATGTTCTCCGCTGGATCGAA 924
Qy 300 ThrThrSerLysAlaThrGluTyrValMetAlaGluLeuMetLysAsnProThrGluLeu 319
Db 925 ACGTATCGACAACATAATTTGGCATGATGATGAAATGATGAAAGAAACAGTGTCTTA 984
Qy 320 ArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGluSer 339
Db 985 GCAAGGCAACAGCTGAAGTGAGACAAGTCTTCAAGGAAAAAGAAAGGTTTTCACAAAT 1044
Qy 340 ArgPheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHisPro 359
Db 1045 GATCTTGTGAGTGAAGTACTTGAAGTTAGTAATCAAGAAACTCTAAGGATGCACTT 1104
Qy 360 ProVal---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyrGlu 378

Db 1105 CCAATTCCTCTTAGTCCCTAGAGAAATGTATGAGGATACAAAGATTGACGGGTACAAAT 1164
Qy 379 IleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsnThr 398
Db 1165 ATACCTTTCAAACTCGAGTCATAGTTAATGATGGCAATTTGGACGAGATCCTGAAAGT 1224
Qy 399 TrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyrLys 418
Db 1225 TGGGATGACCTGAAAGCTTTTCCCAAGAGATTCGAGAAATAGTTCTGTGACCTTCTT 1284
Qy 419 GlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIleThr 438
Db 1285 GGAACCATCATCAATTTATTCATTTGTCGGGAGAGAGATTGCTCTGGAATGCTT 1344
Qy 439 SerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrpGlu 458
Db 1345 TTTGTTTAGCAATGTTGGACAACCATTAGCTCAATTTATCTACTTCGATCGGAAA 1404
Qy 459 LeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyAlaLeu 478
Db 1405 CTCCCTAATGCAAAAGTCACGAAATTTGGACATGACGAGATCACCTGGAATTTCTGCA 1464
Qy 479 ArgLysLysIleAspLeuLysLeuIleProIleProTyr 491
Db 1465 ACAAGAGGATGATCTTGTGTTGATTGCCACCCCTTAT 1503
RESULT 2
US-08-313-075A-29
; Sequence 29, Application US/08313075A
; Patent No. 5639870
; GENERAL INFORMATION:
; APPLICANT: Holton, Timothy A.
; APPLICANT: Cornish, Edwin C.
; APPLICANT: Tanaka, Yoshikazu
; TITLE OF INVENTION: GENETIC SEQUENCES ENCODING FLAVONOID
; NUMBER OF INVENTION: PATHWAY ENZYMES AND USES THEREFOR
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,075A
; FILING DATE: 30-NOV-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PL 1538/92
; FILING DATE: 27-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PL 6698/93
; FILING DATE: 07-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PCT/AU93/00127
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9433
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1665 base pairs

Qy	225	whiStyrllelleGlyAlaGluSerlySProArgArgLeuHlaGlnGluLeuAspIil	245
Db	669	TCATGTACTAAATCGGTATGAAGGGTAAATATTATGGATGTTTCATCGTAAGGTAGTACTGCTAT	728
Qy	245	eLeuGluGluIleLeuAsnGluHis-----LysAlaAsnLysPr	258
Db	729	TGTTGAGGAAGCTATGAATGAGCACAAAGAACTCTTCGAACTGGCAAGACCAATGGTGA	788
Qy	258	oPheGluAlaAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValPr	278
Db	789	AGTGGGAGAGAGAAGATTAAATGATGTATTGCTTAAGACTTAAAGGAAGAGGAGACCTTCA	848
Qy	278	oValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySe	298
Db	849	ACTTCANATCANAATGACACACACTAAAGCCATTTTAAATGACATGTTTCTCGGGGAAC	908
Qy	298	rGluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGl	318
Db	909	AGAAATCTCATCAACAACTAACTGGGCCCATGGTAGAACCTGATGAAAAAATCCACGTGT	968
Qy	318	uLeuArgLysAlaGlnGluGluValArgGlnValPheGlyGluMetGlyLysValAspGl	338
Db	969	ATTCCGGAAGCTCAAGCAGAGGTAAAGAAGTCTTCAAAGGGGAAAGAACTTTCGATGA	1028
Qy	338	uSerArgPheHisAspLeuLysPheLysLeuValValLysGluThrLeuArgLeuHi	358
Db	1029	AGATGATATCGAGGAGCTGAATTACCTTAAGTTAGTCAATTAGAGAAACTTTAAGATCTCA	1088
Qy	358	sProProVal---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTy	377
Db	1089	CCCTCCACTTCCACTTTTCTCTCCAGAGAAGTCTCGGAGAGAAACAGAAATAAATGGCTA	1148
Qy	377	rGluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAs	397
Db	1149	CACATTTCTTTAAATACCAAGTCATAGTTTAATGTTTGGGCTATTTGGAGAGATCCAA	1208
Qy	397	nThrSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTy	417
Db	1209	ATATTGGGATGTCAGAAAGCTTTAAGGCTCGAGAGATTGGAACATAACTCTTTGAATTT	1268
Qy	417	rLysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIil	437
Db	1269	TGCTGCGCAATAATTTTGAATAATCTCTTTGTTGTTGTAAGAGAGATTTCGCCCGAAT	1328
Qy	437	eThrSerAlaIleThrAsnLeuGluTrpValIleIleAsnLeuLeuTrpHisPheAsnTr	457
Db	1329	ATCATTTGGTTTAGCTAATGTTGATCATCATCTGGGCTCAATTTGTTGATCATTTTCGATTG	1388
Qy	457	pGluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyAl	477
Db	1389	GAGACTTCTACTGGGGTGCAGCCAAATGACTTTGAATTGACTAGTTA-GCTGGAGTAAAC	1447
Qy	477	aLeuArgLysLysIleAspLysLeuIleProIleProTrpGlnValSerLeu	495
Db	1448	TACTGGTAGAAAGAGACCTTTACTTGATTTTCACTCTCTTATTCACCTTCTCTA	1502

RESULT 4

US-08-285-309-24

; Sequence 24, Application US/08285309

; Patent No. 5569832

; GENERAL INFORMATION:

; APPLICANT: Holton, Timothy A.

; APPLICANT: Cornish, Edwina C.

; APPLICANT: Kovacic, Filippa

; APPLICANT: Tanaka, Yoshikazu

; APPLICANT: Lester, Diane R.

; TITLE OF INVENTION: GENETIC SEQUENCES ENCODING A 3,5'-

; NUMBER OF INVENTION: HYDROXYLASE AND USES

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Scully, Scott, Murphy & Presser

; STREET: 400 Garden City Plaza

; CITY: Garden City

APPLICANT: Kovacic, Filippa
APPLICANT: Tanaka, Yoshikazu
APPLICANT: Lester, Diane R.
TITLE OF INVENTION: GENETIC SEQUENCES ENCODING A 3,5'-
TITLE OF INVENTION: HYDROXYLASE AND USES
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: U.S.A.
ZIP: 11530

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/502,046
FILING DATE: 14-JUL-1995

CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/285,309

FILING DATE: 03-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.

REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8633Z
TELEPHONE: (516) 742-4343

TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR

INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:

LENGTH: 1666 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-502-046-24

Alignment Scores:
Pred. No.: 2,48e-110 Length: 1666
Score: 1088.00 Matches: 221
Percent Similarity: 62.9% Conservative: 93
Best Local Similarity: 44.3% Mismatches: 168
Query Match: 42.3% Indels: 18
DB: 2 Gaps: 5

US-10-759-813-2 (1-500) x US-08-502-046-24 (1-1666)

QY 12 LeuileSerPheLeuLeuValLeuValVal-MetArgLeuTrpLysLysG1 31
DB 15 TTGGTTCTCTTCCTTATGTAATTTCCCTCATTTTCAAGAAATGGAATC 74
QY 31 nAen-----ProProGlyProTrpLysPheProileGlyAsnLe 46
DB 75 CAATTGTCAACCAAAAAATTCCTCCAGGCCCATCGAAAGTACCTTTCTTGAAGCTT 134
QY 46 uProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGln1 66
DB 135 GCTTCATATGGTA-----GGTGGACTTCCACACCATGCTCTTAGAGATTAGCCAAAA 188
QY 66 eTrGlyProValMetSerLeuGlnileGlyGlnValSerAlaValValIleSerAl 86
DB 189 ATATGACCAATATGACCTTCACTAGTAAATTTCTGCGGTGTAGTTACTTCTCC 248
QY 86 aGluAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIleVa 106
DB 249 TGAGATGGCAAGAAAGTACTAAAAAATCATGACCTTGCAATTTGCATATAGGCTAAACT 308
QY 106 lleuAspAlaGlnileValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGlyAs 126

DB 309 TCTAGGCATTGAGATTGTCTCTATAATAGTTCAGACATTGCCCTTTCCCGTATGGTGA 368
QY 126 pHisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysValG1 146
DB 369 TTACTGGAGGCAAAATGCTAAATTTGTGATTGGAAGTCTTAGTGCCAAAAATGTCG 428
QY 146 nSerSerArgLeuLeuArgGluGluGluMetGluAspAlaIleThrPheLeuArgSerly 166
DB 429 GTCATTTAACTCGATTAGACGAGATGAATACTCTTATGATCGAATTTTTCGCGATCATC 488
QY 166 salaglySerProValIleThrLysIleIleIleIleIleIleIleSerIleMet11 186
DB 489 TCTCGTAAAGCAGTTTATATAACAGAAAGATCTTTTCATTCACAACTCTATGATTG 548
QY 186 eArgThrSerValGly---AenCysLysGlnLysGluArgLeuLeuSerValAlaAspAl 205
DB 549 TAGATCAGTATTTGGGAAAAAGAAATAAGGAGAGAGCAATGTATACGACATGTGAAAAA 608
QY 205 aValAsnGluAlaIleThrSerPheGlyThrAlaAspAlaPheProThrTrpLysLeu 225
DB 609 AATGACAGGCTTAATAGATGGGTTTCGATGTGGCTGACATATTCCTTCGTTGAGGTTCT 668
QY 225 uHisTrpIleIleGlyAlaGluSerLysProArgArgLeuHisGlnIleAspAsp11 245
DB 669 TCATGTACTAATCGGTATGAAGGGTAAATTTATGGATGTTTCATCGTAGGTAGTCTAT 728
QY 245 eLeuGluGluLeuLeuAsnGluHis-----LysAlaAsnLysPr 258
DB 729 TGTGAGGAAGCTATGATGAGCAAAAGAACTCTTCGAACCTGGCAAGACCACTGGTGA 788
QY 258 oPheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValPr 278
DB 789 AGTGGGAGGAGAAAGATTTAATGTATGTTGCTAAGACTTAAAGGAGAGGAGACCTTCA 848
QY 278 oValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySe 298
DB 849 ACTTCCAATCACAATGACACACACTAAGCCATTTTAAATGCATGTTTGTCTCGCGGAAC 908
QY 298 rGluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrG1 318
DB 909 AGAACTTCATCAACAACAATTAACCTGGGCCCATGTAGAACTGTATGAGAAATCCAGTGT 968
QY 318 uLeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspG1 338
DB 969 ATTCCGAAAGCTCAAGCAGAGGTAAAGAGAGTCTTCAAGGGAAGAAACTTTTCGATGA 1028
QY 338 uSerArgPheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHi 358
DB 1029 AGATCATATCGAGGAGCTGAATTACCTTAAGTTAGTATTAGAGAACTTTAAGATCTCA 1088
QY 358 eProProVal---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTy 377
DB 1089 CCCTCCACTCCACTTTGCTTCCAAAGAGAAATGTCGAGAGAAACAGAAATAAATGGCTA 1148
QY 377 rGluIleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAs 397
DB 1149 CACTATTCTTTAAATACCAAGTCATAGTTAAGTTGGGCTATTGGAGAGATCCAAA 1208
QY 397 nThrTrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTy 417
DB 1209 ATATTGGGATGATGCGAGAAAGCTTTAAGCTCGAGAGATTTGAACATAACTCTTTGAATTT 1268
QY 417 rLysGlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGly11 437
DB 1269 TGCTGGCAATAATTTTGAATATCTCTCTTTGTTAGTGGAGAGGATTTGCCCGGAAT 1328
QY 437 eThrSerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTr 457
DB 1329 ATCATTTGGTTAGCTAATGTATCATCCATTCATTCATTCATTCATTCATTCATTCATTC 1388
QY 457 pGluLeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyAl 477

Db	334	AGCATCGGACGAAGATCATGTGTTACGACAACGACGACATCATCTTCAGCCCTACAGC	393
Qy	126	AspHisTrpArgGlnMetLysLysLeuIleTrpIleLeuGluPheLeuSerAlaLysLysVal	145
Db	394	GTGCACCTGGCCCGCCAGATCGGAAAGATCTGGCTCTCCGAGCTCTCAGGCGCCGCAACGTC	453

146 C:\Program Files\Internet Explorer\IEXPLOE.EXE

Qy	146	GlnSerSerArgLeuIleArgGluGluGluMetGluuAspAlaIleInrPheLeuArgSer	161
Db	454	CGCTCTTCGGCTTCATCAGGACGAGGAGGTGCCGCCCTCTCCGGCCACACTCGCGTCC	513
Qy	166	Lys-----AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleSer	183
Db	514	TCGCCCGCGGGGGAGCCGCTGGACCTCAGGAGCGGATAGCCAGCTGAGCTGCTCC	573
Qy	184	IleMetIleArgThrSerValGlyAsn---CysLysGlnLysGluuArgLeuSerVal	202
Db	574	ATCATCTCAGGGCGGCGCTTCGGAGCGGTGATCAGGACACGAGGAGCTGTGTGAGCTG	633
Qy	203	AlaAspAlaValAsnGluuAlaAlaThrSerPheGlyThrAlaAspAlaPheProThrTrp	222
Db	634	GTGAAGAGCGCCCTCAGCATGGCGTCCGGGTTGAGCTCGCGGACATGTTCCCTTCCTCC	693
Qy	223	LysLeuLeuHisTyrIleIleGlyAlaGluSerLysProArgLeuHisGlnGluIle	242
Db	694	AAGCTCTCAACTTGCTCTGCTGGAAACAAGAGCAAGCTGTGGAGGATCGCGCGCGCTC	753
Qy	243	AspAspIleLeuGluGluIleLeuAsnGluHisLysAlaAsnLys-----ProPheGlu	260
Db	754	GACGCCATCTTCGAGGGCCATCGTGGAGGAGCACAAAGCTCAAGAAGAGCGCGGAGTTGGC	813
Qy	261	AlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyValProValPro	280
Db	814	GGCGAGGACATTATTGACGTACTCTTTAGGATGCAGAAGGATAGCCAGATCAAAGTCCCC	873
Qy	281	ValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySerGluThr	300
Db	874	ATCACCAACCAAGGCATCAAAGCCCTTCACTTCGACACGTTCTCAGCGGGGACCGAGACA	933
Qy	301	ThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGluLeuArg	320

DB 934 TCATCAACCAACCACCCIGTGGGTGATGGCGAGCTCAGAGGTGATGGCG 993

Qy	340
Db	105
Qy	360

DB 1054 GTGCAGGAGCTTAAGTACATGAATCGGTGGTGAAGGAGACGATGAGGATGCACCTCCG 1111

[illegible]

Qy 481 LysileAspLeuLysLeuLeuProLeuProTyrGlnValSer 494
 Db 1474 AAGAAATCTTCTACTCGTCCACACCCCTACGATCCTTCC 1515

RESULT 7

US-09-292-768-69
 ; Sequence 69, Application US/09292768
 ; Patent No. 6194185
 ; GENERAL INFORMATION:
 ; APPLICANT: Croteau, Rodney B
 ; APPLICANT: Lupien, Shari L
 ; APPLICANT: Karp, Frank
 ; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR THE PRODUCTION OF
 ; TITLE OF INVENTION: LIMONENE HYDROXYLASES
 ; FILE REFERENCE: wsl13463
 ; CURRENT APPLICATION NUMBER: US/09/292,768
 ; CURRENT FILING DATE: 1999-04-14
 ; EARLIER APPLICATION NUMBER: 08/881,784
 ; EARLIER FILING DATE: 1997-06-24
 ; NUMBER OF SEQ ID NOS: 70
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 69
 ; LENGTH: 1665
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:
 ; OTHER INFORMATION: computer-generated nucleic acid sequence encoding
 ; OTHER INFORMATION: limonene-3-hydroxylase
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)..(1665)
 ; OTHER INFORMATION: computer-generated nucleic acid sequence encoding
 ; OTHER INFORMATION: limonene-3-hydroxylase protein variant
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (19)..(1518)
 ; US-09-292-768-69

Alignment Scores:

Pred. No.: 7,37e-100 Length: 1665
 Score: 993.50 Matches: 205
 Percent Similarity: 61.7% Conservative: 100
 Best Local Similarity: 41.5% Mismatches: 178
 Query Match: 38.6% Indels: 11
 DB: 3 Gaps: 6

US-10-759-813-2 (1-500) x US-09-292-768-69 (1-1665)

Qy 10 SerileLeuileSerPheLeuLeuValLeuileLeuValValMetArgLeuTrpLys 29
 Db 40 GCGATTATAATCTTCTAGTAACCTACACCATATCCTCTAATAATCAAGCAATGGCGA 99
 Qy 30 LysGlnAsnPro-----ProProGlyProTrpLysPheProIleileGlyAsn 45
 Db 100 AAACCGAAACCCCAAGAGAACCTGCTCCGGGCGCGCGAAGCTGCGGTGATCGGGCAC 159
 Qy 46 LeuProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGln 65
 Db 160 CTC---CAC---CTCTATGGGGGAAGCTGCCGACACCGCGCTGGCGAGCTGGCGAAG 213
 Qy 66 IleTyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSer 85
 Db 214 CAGTACGGCCAGTGGCGCAGCGTACAGTCCGCGAGGTGTCTCCGCTGCTGCTCGTCC 273
 Qy 86 AlaGluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIle 105
 Db 274 CGCGAGGCCACGAAGAGGCGATGAAGCTGGTGGACCCCGCTGGCGGACCGGTTCGAG 333
 Qy 106 ValLeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGly 125
 Db 334 AGCATCGGACGAGATCATGTGGTACGACACGACGACATCATCTTCAGCCCTACAGC 393

Qy 126 AspHisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysLysVal 145
 Db 394 GTGCACCTGGCGCCAGATGCGGAAGATCTGCTCCGAGCTCCTCAGCGCCGCGCAACGTC 453
 Qy 146 GlnSerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSer 165
 Db 454 CGTCTCTTCGGCTTCATCAGCAGCAGAGGTGTCCCGCTCCTCGGCCACCTCCGCTCC 513
 Qy 166 Lys-----AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleSer 183
 Db 514 TCGCGCGCGCGGGAGCGCTGAGCTCAGCGAGCGATAGCGACGCTACGCTGCTCC 573
 Qy 184 IleMetIleArgThrSerValGlyAsn---CysLysGlnLysGluArgLeuLeuSerVal 202
 Db 574 ATCATCTGCGGCGCGCTTCGGGAGCGTGTGAGGACCGACGAGGATGCTGTGAGCTG 633
 Qy 203 AlaAspAlaValAsnGluAlaThrSerPheGlyThrAlaAspAlaPheProThrTrp 222
 Db 634 GTGAAGGAGCGCCTCAGCATGCGTCCGGTTCGAGCTCGCGACATGTTCCCTCTCC 693
 Qy 223 LysLeuLeuHisTyrIleIleGlyAlaGluSerLysProArgLeuHisGlnGluLe 242
 Db 694 AAGCTCTCAACTGCTCTGCTGGAACAAGAGCAAGCTGTGAGGATGCGCGCGCTC 753
 Qy 243 AspAspIleLeuGluIleLeuAsnGlnHisLysAlaAsnLys-----ProPheGlu 260
 Db 754 GACGCCATCTCGAGGCCATCGTGGAGCAGACCAAGCTCAAGAGAGCGCGCGAGTTGGC 813
 Qy 261 AlaAspAsnLeuMetAspValLeuLeuAsnGlnLysAsnGlyAsnValProValPro 280
 Db 814 GCGGAGGACATATTGAGCTACTCTTTAGGATGCAGAGGATAGCAGATCAAGATCCC 873
 Qy 281 ValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySerGluThr 300
 Db 874 ATCACCAACCAACGCCCATCAAGCCCTTCATCTCGACACGTTCTCAGCGGGAGCAGACA 933
 Qy 301 ThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGluLeuArg 320
 Db 934 TCATCAACCAACCCCTGTGGTGTGAGGAGCTGTAGGAAATCCAGAGGTGATGGGG 993
 Qy 321 LysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGluSerArg 340
 Db 994 AAAGCGCAGCGGAGTGCAGCGCGCTGAAGGGGAAGAGAGCTGGGAGCTGGACGAC 1053
 Qy 341 PheHisAspLeuLysPhePheLysLeuValLysGluThrLeuArgLeuHisProPro 360
 Db 1054 GTGCAGGAGCTTAAGTACATGAATCGGTGTGTAAGAGAGACGATGAGGATGACCCCTCG 1113
 Qy 361 ValValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyrGluIleHis 380
 Db 1114 ATCCCGTTGATCCCGAGATCATGCGAGAGAAGAAATCGAGGTCAAGGGTACAGATTCCG 1173
 Qy 381 ProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsnThrTrpSer 400
 Db 1174 AATAGGCCAGATCATGATCAACCTGTGCTCCATGGGTAGGAATCTCTCTACTGGAA 1233
 Qy 401 GluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyrLysGlyThr 420
 Db 1234 AAACCCGAGACCTTTTGGCCCGAAAGGTTTGACCAAGTCTCGAGGGAATTCATGGGAAC 1293
 Qy 421 ThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIleThrSerAla 440
 Db 1294 GATTTCCGAGTTTATCCCATTTGGAGCTGGAAGAAGAAATCTGCCCGGTTTGAATTCGGG 1353
 Qy 441 IleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrpGluLeuAla 460
 Db 1354 TTGGCAATGTTGAGTCCCATTTGGCAGACGCTCTTTTACCCTTCGACTCGAAGTGGCG 1413
 Qy 461 AspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyAlaLeuArgLys 480
 Db 1414 GAAGGAATGAACCTTCGATATGACATGTCTGAGGCAAGAGGCTTACCGGAATAAGA 1473

QY 481 LysileAspLeuLysLeuLeuProileProTyrGlnValSer 494
Db 1474 AAGAACAATCTTCTACTCGTTCCACACACCTACGATCCTTCC 1515

RESULT 8

US-09-292-768-5
; Sequence 5, Application US/09292768
; Patent No. 6194185
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Lupien, Shari L
; APPLICANT: Karp, Frank
; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR THE PRODUCTION OF
; FILE REFERENCE: wsu13463
; CURRENT APPLICATION NUMBER: US/09/292,768
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 08/881,784
; EARLIER FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1515
; TYPE: DNA
; ORGANISM: Mentha piperita
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1494)
US-09-292-768-5

Alignment Scores:

Pred. No.: 8,21e-100 Length: 1515
Score: 992.50 Matches: 206
Percent Similarity: 61.5% Conservative: 98
Best Local Similarity: 41.7% Mismatches: 177
Query Match: 38.6% Indels: 13
DB: 3 Gaps: 6

US-10-759-813-2 (1-500) x US-09-292-768-5 (1-1515)

QY 12 LeuileSerPheLeuLeuValLeuLeuValValMetArgLeu----- 27
Db 16 CTTTGTGCGCGCTTATAATCTCTGTAGTAACCTACACCATATCCCTCTTAATCAACCAA 75
QY 28 TrpLysLysGlnAsnPro-----ProProGlyProTrpLysPheProIleIle 43
Db 76 TGGCGAAACCCGAAACCCCAAGGAAAGTTCCCGCGGCGCGCGGAGGCTCGCTGATC 135
QY 44 GlyAsnLeuProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeu 63
Db 136 GGGCACCTC---CAC---CTCCTGTGGGGAGCTCCCGCAGCAGCAGCGCTGGCCAGCGTG 189
QY 64 AlaGlnIleTyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValIle 83
Db 190 GCGAAGGAGTACGCGCCCGCGCGCCAGTGCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 249
QY 84 SerSerAlaGluAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArg 103
Db 250 TCGTCGCGGAGCGACGAGGAGGCGATGAAGCTCGTAGACCCCGCGCGCGCGGACCGG 309
QY 104 ProIleValLeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSer 123
Db 310 TTCGAGAGCATCGGACGAGGATCATGTGTAGCAACAGGAGCATATCTTCAGCCCC 369
QY 124 TyrGlyAspHisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLys 143
Db 370 TACAGGAGCACTGGCGCCGAGATCGCAAGATCGGTCTCCGAGTCTCTCTCTCCCGC 429
QY 144 LysValGlnSerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeu 163
Db 430 AACGTCCTCTTCGCTTCATCCGCGAGGAGGTGTCCGCGCTCCGCGCACCTC 489
QY 164 ArgSerLysAlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleSer 183

Db 490 CGCTCTCGCGGCGCGCGCTGGACATGACGAGGATAGACGCTGACGCTGCTCC 549
QY 184 IleMetIleArgThrSerValGlyAsn---CysLysGlnLysGluArgLeuLeuSerVal 202
Db 550 ATCATCTGCGGCGCGCTTCCGAGCGTATCAGGCAACACGCGAGCTGTGGGCTG 609
QY 203 AlaAspAlaValAsnGluAlaThrSerPheGlyThrAlaAspAlaPheProThrTrp 222
Db 610 GTCAGGACGCGCTCAGCATGCGCTCGGGTTCGAGCTCCCGACATGTTCCCTCTCC 669
QY 223 LysLeuLeuHisTyrIleIleGlyAlaGluSerLysProArgLeuGlnHisGlnIle 242
Db 670 AAGCTCTCAACCTCTCTGCTGAACAAGCAAGCTCTGGAGGATGCCCGCGGCTC 729
QY 243 AspAspIleLeuGluGluLeuLeuAsnGluHisLysAlaAsnLys-----PropheGlu 260
Db 730 GACACCATCTCGAGGCCATCTGTCGACGACCAAGTTCAGGAAGAGCGCGAGTTCGGC 789
QY 261 AlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValProValPro 280
Db 790 GCGGAGGACATCATGACGCTCTCTTCAGGATGCAAGAGCCACCCAGATCAAGTCCCC 849
QY 281 ValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySerGluThr 300
Db 850 ATCACCACCACTCCATCAAGCTTCATCTTCGATACGTTCTCAGCAGGAGTGTGACA 909
QY 301 ThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGluLeuArg 320
Db 910 TCCTCAACCAACCCCTATGCGGTGCTGGCGGAGCTGATGAGGAACCCGCGAGTGTGGC 969
QY 321 LysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGluSerArg 340
Db 970 AAAGCGCAGCGGAGGTGAGCGGCACTGAAGAGAAAGACGAACTGGGACGTGGATAT 1029
QY 341 PheHisAspLeuLysPheLysLeuValValLysGluThrLysLeuArgLeuHisProPro 360
Db 1030 GTGCAAGAGCTTAAGTACATGAAATCGGTGGTGAAGGACGATGAGGATGACCTCCG 1089
QY 361 ValValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyrGluIleHis 380
Db 1090 ATCCCGTGTATCCGAGATCATGCAAGAAAGTTCGTTAAACGGGTATACGATTCCG 1149
QY 381 ProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsnThrTrpSer 400
Db 1150 AACAGGCGCAGATCATGATCAACGCTGTGTCATGGCGAGGAATCTCTTACTGGGAA 1209
QY 401 GluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyrLysGlyThr 420
Db 1210 AACCCGATACCTTTTGGCCCGAAGTTTGACCAAGTTTCAAGGATTTTCATGGGAAAT 1269
QY 421 ThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIleThrSerAla 440
Db 1270 GATTTCGAGTTCGTCGCGTTCGAGCGGGAAGAAGTCTGCCCGCTTGAATTCGGG 1329
QY 441 IleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrpGluLeuAla 460
Db 1330 TCGGCAACGTTGAGTTCATTTGGCGCAGCTTCTTACCACTTCGACTCGAAGTGGCG 1389
QY 461 AspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAlaLeuArgLys 480
Db 1390 GAAGGATGAACCTTCTGATATGACATGCTCGAGCGGAGGCTTACCGGAATACTA 1449
QY 481 LysIleAspLeuLeuIleProIleProTyrGlnValSer 494
Db 1450 AAGAACAATCTTCTTCTTGTTCCTCCACACCTACGATCCTTCC 1491

RESULT 9

US-08-881-784-8
; Sequence 8, Application US/08881784
; Patent No. 6083731
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B.

APPLICANT: Lupien, Shari L.
APPLICANT: Karp, Frank
TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR
THE PRODUCTION OF LIMONENE HYDROXYLASES
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSES: Christensen, O'Connor, Johnson and Kindness
ADDRESS: PLLC
STREET: 1420 Fifth Avenue, Suite 2800
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: US/08/881,784

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Shelton, Dennis K.
REGISTRATION NUMBER: 26,997
REFERENCE/DOCKET NUMBER: WSUR19777
TELEPHONE: (206) 224-0718
TELEFAX: (206) 224-0779
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1665 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Mentha x piperita
IMMEDIATE SOURCE:
CLONE: pPWL7
US-08-881-784-8

Alignment Scores:
Pred. No.: 9.51e-100 Length: 1665
Score: 992.50 Matches: 205
Percent Similarity: 61.5% Conservative: 99
Best Local Similarity: 41.5% Mismatches: 179
Query Match: 38.6% Indels: 11
DB: 6 Gaps: 6

US-10-759-813-2 (1-500) x US-08-881-784-8 (1-1665)

Qy 10 SerIleLeuIleSerPheLeuLeuValValMetArgLeuTrpLys 29
Db 40 GCATTATATCTTCTAGTAACCTACACATATCTCTCTATATCAAGCAATGGCGA 99
Qy 30 LysGlnAsnPro-----ProProGlyProTrpLysPheProIleLeuGln 45
Db 100 AACCCGAAACCCCAAGAACCTGCTCCGGGCGCCGCGAAGTGGCGTGTATCGGGCAC 159
Qy 46 LeuProHisLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuGln 65
Db 160 CTC---CAC---CTCCTATGGGGAGAGCTGCCGACGACGCGCTGGCCAGCGTGGCGAAG 213
Qy 66 IleTyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSer 85
Db 214 CAGTAGCGCCAGCGGCGGCGACGTGCGAGCTCGGCGAGGTGTCTCCGTCTGCTCTCTCTC 273
Qy 86 AlaGluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIle 105
Db 274 CGCGAGGCCACGAAGGCGGATGAAGCTGGTGGACCCCGGCTCGCGGACCGGTTCCGAG 333
Qy 106 ValLeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGly 125

Db 334 AGCATCGGGACGAAGATCATGTGTACGACAAACGACGACATCATCTTCAGCCCTACAGC 393
Qy 126 AspHisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLysVal 145
Db 394 GTGCACCTGGCGCCAGATGCGGAAGATCTGCTCCGAGCTCTCAGCGCCGCAACGTC 453
Qy 146 GlnSerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeuArgSer 165
Db 454 CGCTCTTCGGCTTCATCAGGACGAGAGTGTCGCCCTCCCTCGGCACCTCCGCTCC 513
Qy 166 Lys-----AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleSer 183
Db 514 TCGGCGCGCGGGGAGCGCTGACCTCAGGAGCGGATAGCAGCTGACGTGCTCC 573
Qy 184 IleMetIleArgThrSerValGlyAsn---CysLysGlnLysGluArgLeuLeuSerVal 202
Db 574 ATCATCTGCGCGCGCTTCGGGAGCGTGTAGGAGCCAGGACGAGGAGCTGTGGAGCTG 633
Qy 203 AlaAspAlaValAsnGluAlaThrSerPheGlyThrAlaAspAlaPheProThrTrp 222
Db 634 GTGAAGGACGCCCTCAGCATGCGGCTCGAGCTCGCGGACATGTTCCTCCCTCC 693
Qy 223 LysLeuLeuHisTyrIleIleGlyAlaGluSerLysProArgArgLeuHisGlnGluIle 242
Db 694 AAGCTCTCAACTGCTCTGCTGGNACAAGAGCAAGCTGTGGAGGATGCGCGCGCTC 753
Qy 243 AspAspIleLeuGluGluIleLeuAsnGluHisLysAlaAsnLys-----ProPheGlu 260
Db 754 GACGCCATCTCGAGGCCATCGTGGAGGAGCACAAAGCTCAAGAGAGCGCGGAGTTGCG 813
Qy 261 AlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValProValPro 280
Db 814 GCGGAGGACATTATTGACGTACTCTTTAGGATGCAAGAGGATAGCCAGATCAAGATCCC 873
Qy 281 ValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySerGluThr 300
Db 874 ATCACCACCAACGCCATCAAGCCCTTCATCTTCGACAGCTTCTCAGCGGGAGCCAGACA 933
Qy 301 ThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGluLeuArg 320
Db 934 TCATCAACCAACCCCTGTGGGTGATGCGGAGCTGATGAGGAATCCAGAGGTGATGGCG 993
Qy 321 LysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGluSerArg 340
Db 994 AAAGCGCAGCGGAGGTGAGCGCGCTGAAGGGGAGAGACGCGACTGGGACGTGGACGAC 1053
Qy 341 PheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHisProPro 360
Db 1054 GTGCAGGAGCTTAAGTACATGAATCGGTGGTGAAGGAGCAGTGAAGGATGACCCCTCG 1113
Qy 361 ValValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyrGluIleHis 380
Db 1114 ATCCCGTTGATCCCGAGATCATGCAGAGAAGAATGCGAGTCAACGGGTACACGATCCG 1173
Qy 381 ProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsnThrTrpSer 400
Db 1174 AATAAGGCCAGAAATCATGATCAACGTGGTCCATGGGTAGGAATCTCTCTACTGGGA 1233
Qy 401 GluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyrLysGlyThr 420
Db 1234 AAACCGAGACCTTTTGGCCGAAAGTTTGACCAAGTCTCGAGGATTCATGGGAAC 1293
Qy 421 ThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIleThrSerAla 440
Db 1294 GATTTTCAGTTTCCATCCCATTTGGAGCTGGAAGAAGAAATCTGCCCGGTTGAAATTCGGG 1353
Qy 441 IleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrpGluLeuAla 460
Db 1354 TTGGCAATGTTGAGGTCCCATTTGGCAGCTCTTTTACACTTCGACTCGGAAGTTGGCG 1413
Qy 461 AspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyAlaLeuArgLys 480

Db 1414 GAAGGAATGAACCTTTCCGATATGGACATGTCTGAGGCGAAGGCTTACCGGAATAAGA 1473
QY 481 LysileAspLeuLysLeuLeuProIleProTyrGlnValSer 494
Db 1474 AAGACAATCTTCTACTGTTCCACACACCTCAGATCCTTCC 1515

RESULT 10

US-09-292-768-3
; Sequence 3, Application US/09292768
; Patent No. 6194185
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Lupien, Shari L
; APPLICANT: Karp, Frank
; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR THE PRODUCTION OF
; FILE OF INVENTION: LIMONENE HYDROXYLASES
; CURRENT APPLICATION NUMBER: US/09/292,768
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 08/881,784
; EARLIER FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: Mentha piperita
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19) .. (1518)
US-09-292-768-3

Alignment Scores:

Pred. No.: 9,51e-100 Length: 1665
Score: 92.50 Matches: 205
Percent Similarity: 61.5% Conservative: 99
Best Local Similarity: 41.5% Mismatches: 179
Query Match: 38.6% Indels: 11
DB: 3 Gaps: 6

US-10-759-813-2 (1-500) x US-09-292-768-3 (1-1665)

QY 10 SerIleLeuLeuSerPheLeuValLeuValLeuValMetArgLeuTrpLys 29
Db 40 GGGATTATAATCTTGATGTAACCTACACCATATCCTCTTAATCAAGCAATGGCGA 99
QY 30 LysGlnAsnPro-----ProProGlyProTrpTrpLysPheProIleLeuGlyAsn 45
Db 100 AAACCGAAACCCCAAGAGAACCTGCTCCGGGCGCCGCGAAGCTGCGCTGATCGGGCAC 159
QY 46 LeuProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeuAlaGln 65
Db 160 CTC---CAC---CTCTATGGGGGAAGCTGCCGAGCAGCGCTGGCCAGCGTGGCGAAG 213
QY 66 IleTyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIleSerSer 85
Db 214 CAGTAGCGGCCAGTGGCGACGCTGAGCTGCGGAGGTGTCTCGTGTGCTCGTCTCGTCC 273
QY 86 AlaGluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArgProIle 105
Db 274 CGCGAGGCCACGAAGGAGCGGATGAAGCTGGTGGACCCCGCTGGCGGACCGGTTTCGAG 333
QY 106 ValLeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSerTyrGly 125
Db 334 AGCATCGGGACGAAGATCATGTGTACGACAAACGACGACATCATCTTCAGCCCCCTACAGC 393
QY 126 AspHisTrpArgGlnMetLysIleTrpIleLeuGluPheLeuSerAlaLysLysVal 145
Db 394 GTGCACTGGCGGCAGATGCGGAAGATCTCGCTCTCGAGCTCTCAGCGCCCGCAACGTC 453
QY 146 GlnSerSerArgLeuIleArgGluGluGluMetGluAspAlaIleThrPheLeuArgSer 165
Db 454 CGCTCTCTCGGCTTTCATCAGGCGAGGAGGTGTCCCGCTCTCTCGGCCACCTCGGCTCC 513

RESULT 11

US-08-881-784-5
; Sequence 5, Application US/08881784
; Patent No. 6083731

QY 166 Lys-----AlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleIleSer 183
Db 514 TCGCGCGCGCGGGGAGCGCGTGGACCTCAGCGCGGATAGCGACGCTGCGTGTCTCC 573
QY 184 IleMetIleArgThrSerValGlyAsn---CysLysGlnLysGluLysGluArgLeuSerVal 202
Db 574 ATCATCTGCAGGCGCGCTTCGGAGCGTGATCAGGACCCACGAGGAGCTGGTGGAGCTG 633
QY 203 AlaAspAlaValAsnGluAlaThrSerPheGlyThrAlaAspAlaPheProThrTrp 222
Db 634 GTGAAGGACGCCCTCAGCATGGCTCCGGGTTCGAGCTCGCCGACATGTTCTCCCTCCTCC 693
QY 223 LysLeuLeuHisTyrIleIleGlyAlaGluSerLysProArgArgLeuHisGlnGluIle 242
Db 694 AAGCTCTCTCACTTGTCTGCGAAACAGAGCAAGCTGTGGAGGATGCCGCCCGCTCC 753
QY 243 AspAspIleLeuGluIleLeuAsnGluHisLysAlaAsnLys-----ProPheGlu 260
Db 754 GACGCCATCTCGAGGCCATCTGGAGGAGCACAGCTCAAGAGAGCGCGGAGTTTGGC 813
QY 261 AlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValProValPro 280
Db 814 GCGGAGGACATTATTGACGTACTCTTTAGGATGCAGAGGATAGCCAGATCAAGATCCCC 873
QY 281 ValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySerGluThr 300
Db 874 ATCACCACCAAGCCATCAAGCCCTTCTTCGACACCGTCTTCAGCGGGGACCGAGACA 933
QY 301 ThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGluLeuArg 320
Db 934 TCATCAACCCACCCTCTGGGTGTGGCGGAGCTGATCAGGAATCAGAGGTGTATGGCG 993
QY 321 LysAlaGlnGluGluValArgGlnValPheGlyGluMetGlyLysValAspGluSerArg 340
Db 994 AAAGCGCAGCGCGAGGTGAGAGCGCGCTGAAGGGGGAAGACGAGCTGGGACGTGGACGAC 1053
QY 341 PheHisAspLeuLysPhePheLysLeuValLysGluThrLeuArgLeuHisProPro 360
Db 1054 GTGCAGGAGCTTAAGTACATGAATTCGGTGGTGAAGGAGCAGTGAAGATGACCCCTCCG 1113
QY 361 ValValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyrGluIleHis 380
Db 1114 ATCCCGTTGATCCCGAGATCATGCAGAGAAGAATCGAGGTCAACGGGTACACGATTCCG 1173
QY 381 ProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsnThrTrpSer 400
Db 1174 AATAAGGCCAGAATCATGATCAACGTGTGTCTCATGGGTAGGAATCTCTCTACTGGGAA 1233
QY 401 GluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyrLysGlyThr 420
Db 1234 AAACCCGAGACCTTTTGGCCCGAAGTTTGAACCAAGTCTCGAGGATTTTCATGGGAAC 1293
QY 421 ThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIleThrSerAla 440
Db 1294 GATTTCCGAGTTCCATCCATTTGGAGCTGGAAGAAGAATCTGCCCGTTTGAATTTCCGG 1353
QY 441 IleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrpGluLeuAla 460
Db 1354 TTGGCAAAATGTTGAGGTCCCATTTGGCAGACGCTTCTTTACCACTTCGACTGGAAGTTGGG 1413
QY 461 AspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAlaLeuArgLys 480
Db 1414 GAAGGAATGAACCTTCCGATATGGACATGTCTGAGGCGAAGAGGCTTACCGGAATAAGA 1473
QY 481 LysIleAspLeuLysLeuIleProIleProTyrGlnValSer 494
Db 1474 AAGACAATCTTCTACTGTTCCACACACCTCAGATCCTTCC 1515

;; GENERAL INFORMATION:
;; APPLICANT: Groteau, Rodney B.
;; APPLICANT: Lupien, Shari L.
;; APPLICANT: Karp, Frank
;; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR
;; TITLE OF INVENTION: THE PRODUCTION OF LIMONENE HYDROXYLASES
;; NUMBER OF SEQUENCES: 58
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Christensen, O'Connor, Johnson and Kindness
;; ADDRESS: PLIC
;; STREET: 1420 Fifth Avenue, Suite 2800
;; CITY: Seattle
;; STATE: WA
;; COUNTRY: USA
;; ZIP: 98101

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: IBM PC compatible
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/881,784

;; FILING DATE:
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Shelton, Dennis K.
;; REGISTRATION NUMBER: 26,997
;; REFERENCE/DOCKET NUMBER: WSUR19777
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (206) 224-0718
;; TELEFAX: (206) 224-0779

;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1762 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; ORIGINAL SOURCE:
;; ORGANISM: Mentha spicata
;; INDIVIDUAL ISOLATE: cDNA encoding
;; IMMEDIATE SOURCE:
;; CLONE: pSM12.2

;; NAME/KEY: misc feature
;; LOCATION: 558..1212
;; OTHER INFORMATION: /product= "Probe LH-1 (Figure 4A)"
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: 39..538
;; OTHER INFORMATION: /product= "Probe LH-2 (Figure 4A)"
;; US-08-881-784-5

Alignment Scores:
Pred. No.: 1.03e-98 Length: 1762
Score: 983.50 Matches: 199
Percent Similarity: 62.1% Conservative: 108
Best Local Similarity: 40.3% Mismatches: 170
Query Match: 38.2% Indels: 17
DB: 3 Gaps: 7

US-10-759-813-2 (1-500) x US-08-881-784-5 (1-1762)

QY 12 LeuileSerPheLeuValLeuValLeuValValMetArgLeu----- 27
DB 32 CTTTGTGGCAATATATATCTTGTGGCAACTACATCGTATCCCTCTTAATCAACCAA 91

QY 28 TrpLysLys-----GlnAsnProProGlyProTrpLysPheProLeu 43
DB 92 TGGCGAAATCGAAATCCCAACAAACCTACCTCCGAGCCCTCCGAAAGTCCGCGTATC 151

QY 44 GlyAsnLeuProHisLeuLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeu 63

Db 152 GCCCACCTC---CAC---TTCTGTGGGAGGGCTTCCCCAGCAGCTGTTTAGGAGCATA 205

QY 64 AlaGlnIleTrpGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIle 83
Db 206 GCCCAGAAGTACGGCCGGTGGCCACGTGCGAGTGGGAGAAGTACTACTCGTGTGCTG 265

QY 84 SerSerAlaGluAlaAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArg 103
Db 266 TCGTCGGCGAGGAGCGAAGCGGATGAAGGTGCTGGACCCCGAACTTCGCCACCGG 325

QY 104 ProIleValLeuAspAlaGlnIleValPheTrpAsnArgLysAspValPheAlaSer 123
Db 326 TTCGAGCGCATCGGGTCCAGGACCATGGGTACGACAAGATGATCATCTTCAGCCCT 385

QY 124 TyrGlyAspHisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLys 143
Db 386 TACAACGATCACTGGCGCCAGATCGGAGGATCTCGGTGACAGAGCTGCTGAGCCGAAG 445

QY 144 LysValGlnSerSerArgLeuIleArgGluGluGlnMetGluAspAlaIleThrPheLeu 163
Db 446 AACGTCAGTCTTCGGGTACATAAGGCAGGAGGATCGAGCGCTCATCGCGTGTCTC 505

QY 164 ArgSerLysAlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleSer 183
Db 506 GGGTCGTCGGGGGAGCCCGTCACTGACGAGGAGGTGTCGAAGATGCTGCTGTCTC 565

QY 184 IleMetIleArgThrSerValGlyAsnCysLysGlnLysGluArgLeuLeuSerValAla 203
Db 566 GTCTGTGTCAGGGCGGCTTCGGGAGTGTCTCAAGGACACGGGT-----TCGTTGGCG 619

QY 204 AspAlaValAsnGluAla-----AlaThrSerPheGlyThrAlaAspAlaPhePro 220
Db 620 GAGTTGGTGAAGGAGTCTCGCTCGCATTTGGCGTTCGGGTGTGAGCTGGCGGATCTCACCCT 679

QY 221 ThrTrpLysLeuLeuHisTrpIleGlyAlaGluSerLysProArgArgLeuHisGln 240
Db 680 TCCTCATGGCTCTCAACTGCTTAGTTGAACAAGTACAGGTTGCAGAGGATGCGCCGC 739

QY 241 GluIleAspAspIleLeuGluGluIleLeuAsnGluHisLysAlaAsnLys-----Pro 258
Db 740 CGCTCGATCATCTTGTATGGTTCCTGGAGGAGCATAGGAGAGAAGACGCGCGAG 799

QY 259 PheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuLysAsnGlyAsnValPro 278
Db 800 TTTGAGGCGAGGACATCGTCGAGTCTTTTCAGATGCAGAGGCGAGCCACATCAAA 859

QY 279 ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer 298
Db 860 ATTCCCATTAATTCGAATTCGATCAAGGGTTTCATTTTCGACACCTTCTCCGCGGAGCT 919

QY 299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
Db 920 GAAACGCTCTCGACGACCATCTCATGGGCGTGTTCGGAACATGATGAGGAATCCGCGAAG 979

QY 319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
Db 980 ATGGCCAAAGTTCAGGCGGAGTAAAGAGGCGCTCAAGGGGAAAGACAGCTCGTGGATTG 1039

QY 339 SerArgPheHisAspLeuLysPhePheLysLeuValLysGluThrLeuArgLeuHis 358
Db 1040 AGCGAGGTGCAAGAGCTAAATACCTGAGATCGGTGTTAAAGGAGACTCTGAGGCTGCAC 1099

QY 359 ProProValValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTrpGlu 378
Db 1100 CTTCCCTTTCCATTAATCCCAAGACAATCCAGGGAAGATGCGAGGTAAACGGGTACACG 1159

QY 379 IleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsnThr 398
Db 1160 ATTCCGCGCAAACTAGAAATCTTCATCAACGTCTGGCTATCGAAGGATCCCAATAC 1219

QY 399 TrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTrpLys 418

```
Db 1220 TGGGAAGATCCGACACCTTCGCCCTGAGAGATTCGATGAGGTTTCCAGGATTTTCATG 1279
QY 419 GlyThrThrPheGluLeuValProPheGlyValGlyLeuArgLeuGlyLeuThr 438
Db 1280 GGAACAGATTTCAGTTTCATCCATTCGGGGCGGTTCGAAGAATCTGCCCGGTTCATAC 1339
QY 439 SerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrpGlu 458
Db 1340 TTCGGGCTGGCAAAATGTTGAGATCCCATTCGGCGCAACTGCTCTACCACTTCGACTGGAA 1399
QY 459 LeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAlaLeu 478
Db 1400 TTCGCCAAGGAATGACTGATGCCGACTTGGACATGACGAGACCCCGAGGTCTTCTGGG 1459
QY 479 ArgLysLysIleAspLeuLysLeuIleProIleProTyrGln 492
Db 1460 CCAAAAAGAAAATGTTTGGTTGGTCCACACTCTATAAA 1501

RESULT 12
US-09-292-768-1
; Sequence 1, Application US/09292768
; Patent No. 6194185
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Lupien, Shari L
; APPLICANT: Karp, Frank
; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR THE PRODUCTION OF
; TITLE OF INVENTION: LIMONENE HYDROXYLASES
; FILE REFERENCE: waur13463
; CURRENT APPLICATION NUMBER: US/09/292,768
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 08/881,784
; EARLIER FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1762
; TYPE: DNA
; ORGANISM: Mentha spicata
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (20)..(1507)
US-09-292-768-1

Alignment Scores:
Pred. No.: 1,03e-98 Length: 1762
Score: 983.50 Matches: 199
Percent Similarity: 62.1% Conservative: 108
Best Local Similarity: 40.3% Mismatches: 170
Query Match: 38.2% Indels: 17
DB: 3 Gaps: 7

US-10-759-813-2 (1-500) x US-09-292-768-1 (1-1762)
QY 12 LeuIleSerPheLeuLeuValLeuIleLeuValValMetArgLeu----- 27
Db 32 CTTTTCGGCAATTATATCTTGTGGCAACCTACATCGTATCCCTCTCAACACAA 91
QY 28 TrpLysLys-----GlnAsnProProGlyProGlyProTrpLysPheProIleIle 43
Db 92 TGGCGAAAATCGAAAATCCCAACAAAACCTACCTCCGAGCGCTCCGAAAGCTGCCGGTGTATC 151
QY 44 GlyAsnLeuProHisLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeu 63
Db 152 GGCACCTC---CAC---TTCCTGTGGGAGGCTTCCCAGCAGCAGCTGTTTAGGACATA 205
QY 64 AlaGlnIleTyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValValIle 83
Db 206 GCCCAGAAGTACGGGCGGTGGCGCAGTGCAGCTGGGAGAAGTGTACTCGGTGGTGTCTG 265
QY 84 SerSerAlaGluAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArg 103
Db 266 TCGTCGGCGGAGCGACGGAAGCAGGCGATGAAGGTCTGGACCCCGAACTTCGCCGACCGG 325
```

```
QY 104 ProIleValLeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSer 123
Db 326 TTCAGCGGATCCGGTCCAGGACCATGTGTGATACGACAAAGATGACATCATCTTCAGCCCT 385
QY 124 TyrGlyAspHisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLys 143
Db 386 TACAACGATCACTGGCGCCAGATCGCGAGGATCTCGGTGACACAGCTGCTGAGCCCGAAG 445
QY 144 LysValGlnSerSerArgLeuIleArgGluGluGluMetGluAspAlaIleThrPheLeu 163
Db 446 AACGTGAGTCTCTCCGGTACATAAGCGAGGAGAGATCGAGCGCTCATCCGCTGCTC 505
QY 164 ArgSerLysAlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleSer 183
Db 506 GGGTCGTGGGGGAGCGCGGTGACGTGACGAGGAGGTGTCGAAGATGTCGTGTGTC 565
QY 184 IleMetIleArgThrSerValGlyAsnLysLysGlnLysGluArgLeuLeuSerValAla 203
Db 566 GTCGTGTGAGGCGCGGTTTCGGAGTGTCTCAAGGACCAGGCT-----TCGTTGGCG 619
QY 204 AspAlaValAsnGluAla-----AlaThrSerPheGlyThrAlaAspAlaPhePro 220
Db 620 GAGTTGGTGAAGGAGTCTGCTGGCATTTGGCGTCCGGGTTTCAGCTGGCGGATCTCTACCT 679
QY 221 ThrTrpLysLeuLeuHisTyrIleIleGlyAlaGluSerLysProArgArgLeuHisGln 240
Db 680 TCTCATGGCTCTCAACCTGTAGCTTGAACAAGTACAGGTTGACAGGATGCGCGCG 739
QY 241 GluIleAspAspIleLeuGluIleLeuAsnGluHisLysAlaAsnLys-----Pro 258
Db 740 CGCCTCGATCACATCCTTCATGGTTCCTGGAGGAGCATAGGAGAGAAGAAGCGCGGAG 799
QY 259 PheGluAlaAspAsnLeuMetAspValLeuLeuAsnLeuGlnLysAsnGlyAsnValPro 278
Db 800 TTTGGAGCGGAGGACATCGTCGACGTTCTTTTCAGGATCGAAGAGGCGAGCATCAAA 859
QY 279 ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer 298
Db 860 ATTCCATTACTTCCAAATGTCATCAAGGTTTCATTTTCGACACCTTCTCCGCGGAGCT 919
QY 299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
Db 920 GAAACGTCTTCGACGACCATCTCATGGCGGTTGTCGGAATCTGATGAGGAATCCGCGAAG 979
QY 319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
Db 980 ATGGCCCAAGGTGCGCGGAGGTAAAGAGAGGCGCTCAAGGGAAGACAGTCTGTGGATTG 1039
QY 339 SerArgPheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHis 358
Db 1040 ACGGAGGTGCAAGAGCTAAATACCTGAGATCGGTGTTTAAAGAGAGACTCTGAGGCTGCGAC 1099
QY 359 ProProValValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyrGlu 378
Db 1100 CTTCCCTTTCATTAATCCAGACAATCCAGGAGAAGTCCGAGGTAAACGGGTACAG 1159
QY 379 IleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsnThr 398
Db 1160 ATTCCGGCCAAAACATAGAAATCTTCATCAACGCTCGGGCTATCGGAAGGATCCCCAATAC 1219
QY 399 TrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyrLys 418
Db 1220 TGGGAAGATCCCGACACCTTCGCCCTGAGAGATTCGATGAGGTTTCCAGGATTTTCATG 1279
QY 419 GlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIleThr 438
Db 1280 GGAACGATTTTCAGTTTCATCCATTCGGGGCGGTTCGAAGAATCTGCCCGGTTTACAT 1339
QY 439 SerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrpGlu 458
Db 1340 TTCGGGCTGGCAAAATGTTGAGATCCCATTTGGCGCAACTGCTCTACCACTTCGACTGGAA 1399
```


Db 1400 TTGCCAAGAATGACTGTCGCGACTTGGACATGACGAGACCCCGAGTCTTCTCTGGG 1459
Qy 479 ArgLysLysIleAspLeuLysLeuIleProIleProTyrGln 492
Db 1460 CCAAAAAGAAAATGTTGCTTGGTTCCACACACTCTATAAA 1501

RESULT 14

US-09-292-768-63
; Sequence 63, Application US/09292768
; Patent No. 6194185
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Lupien, Shari L
; APPLICANT: Karp, Frank
; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR THE PRODUCTION OF
; FILE REFERENCE: waur13463
; CURRENT APPLICATION NUMBER: US/09/292,768
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 08/881,784
; EARLIER FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 63
; LENGTH: 1762
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: computer-generated nucleic acid sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (20)..(1507)
; OTHER INFORMATION: Computer-generated nucleic acid sequence encoding
; OTHER INFORMATION: limonene-6-hydroxylase variant
US-09-292-768-63

Alignment Scores:
Pred. No.: 3-7e-98 Length: 1762
Score: 978.50 Matches: 199
Percent Similarity: 61.9% Conservative: 107
Best Local Similarity: 40.3% Mismatches: 171
Query Match: 38.0% Indels: 17
DB: 3 Gaps: 7
US-10-759-813-2 (1-500) x US-09-292-768-63 (1-1762)

Qy 12 LeuIleSerPheLeuLeuValLeuIleLeuValValValMetArgLeu----- 27
Db 32 CTTTGTGCGCAATTATATCTTGTGGCAACTACATCGTATCCCTCTCTAATCAACCAA 91
Qy 28 TrpLysLys-----GlnAsnProProGlyProTrpLysPheProIleIle 43
Db 92 TGGCGAAAATCGAAATCCCAACAAACCTACCTCCGAGCCCTCCGAAGCTGCGGTGATC 151
Qy 44 GlyAsnLeuProHisLeuLeuThrSerAspLeuGlyHisGluArgPheArgAlaLeu 63
Db 152 GGCCACCTC---CAC---TTCTGTGGGGGGGCTTCCCCAGCAGCAGCTGTTAGGACATA 205
Qy 64 AlaGlnIleTyrGlyProValMetSerLeuGlnIleGlyGlnValSerAlaValIle 83
Db 206 GCCCAGAAGTACGGGCGGTGGCGCAGTGCAGCTGGGAGAGTGTACTCGGTGGTGTGCTG 265
Qy 84 SerSerAlaGlnAlaLysGluValMetLysThrGlnAlaAspAlaPheAlaGlnArg 103
Db 266 TCGTCGGCGAGGACGACGAGGAGCGATGAAGTGCTGGACCCCGAACTTCGCCGACCGG 325
Qy 104 ProIleValLeuAspAlaGlnIleValPheTyrAsnArgLysAspValLeuPheAlaSer 123
Db 326 TTCGACGGCATCGGTTCAGGACCAATGTGTACGACAAAGATGACATCTTCAGCCCT 385
Qy 124 TyrGlyAspHisTrpArgGlnMetLysLysIleTrpIleLeuGluPheLeuSerAlaLys 143

Db 386 TACAACGATCACTGGCGCCAGATCGCGAGGATCTGGTGACAGAGCTGTGAGCCCAAG 445
Qy 144 LysValGlnSerSerArgLeuIleArgGluGluMetGluAspAlaIleThrPheLeu 163
Db 446 AACGTGAGTCTCTTCCGGGTACATAAGGACGAGGAGATCGAGGCGCTCATCCGGCTGTC 505
Qy 164 ArgSerLysAlaGlySerProValAsnIleThrLysIleIleTyrGlyIleIleSer 183
Db 506 GGGTCTCGGGGGAGCGCCGTCGACGTGACGAGGAGGTGTCGAAGATGTCTGTGTGTC 565
Qy 184 IleMetIleArgThrSerValGlyAsnCysLysGlnLysGluArgLeuLeuSerValAla 203
Db 566 GTCGTGTGACGCGCGCGTTCGGGAGTGTCTCAAGGACCCAGGT-----TCGTTCGCG 619
Qy 204 AspAlaValAsnGluAla-----AlaThrSerPheGlyThrAlaAspAlaPhePro 220
Db 620 GAGTGTGTAAGGAGTCTGCTGGCATTTGGCGTTCGGGTTCGAGCTGGCGGATCTCTACCT 679
Qy 221 ThrTrpLysLeuLeuHisTyrIleIleGlyAlaGluSerLysProArgArgLeuHisGln 240
Db 680 TCCTCATGGCTCTCAACCTGCTAGCTTGAACAAGTACAGTTGTCAGAGGATGCGCGCG 739
Qy 241 GluIleAspAspIleLeuGluIleLeuAsnGluHisLysAlaAsnLys-----Pro 258
Db 740 CGCTCGATCACATCTTGATGGGTTCTGGAGGACATAGGAGAGAGAGCGCGCGAG 799
Qy 259 PheGluAlaAspAsnLeuMetAspValLeuLeuAsnGlnLysAsnGlyAsnValPro 278
Db 800 TTTGGAGGCGGACATCGTCGACGTTCTTTTTCAGGATGCAAGAGGACGACGACATCAA 859
Qy 279 ValProValThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySer 298
Db 860 ATTCCTTACTTCCAAATTCATCAAGGTTTCATTTTCGACACCTTCTCCGCGGAGCT 919
Qy 299 GluThrThrSerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGlu 318
Db 920 GAAACGTCTTCGACGACCATCTCATGGCGGTTGTCGGAATGATGAGGAATCCGCGCAAG 979
Qy 319 LeuArgLysAlaGlnGluValArgGlnValPheGlyGluMetGlyLysValAspGlu 338
Db 980 ATGGCCCAAGTGCAGGCGAGGTAAAGAGAGCGCTCAAGGGGAAAGACAGTCGTGGATTG 1039
Qy 339 SerArgPheHisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHis 358
Db 1040 ACGGAGTCAAGAGCTAAATACCTGAGATCGTGTGTTAAGAGAGACTCTGAGGCTGCAC 1099
Qy 359 ProProValValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTyrGlu 378
Db 1100 CTTCCCTTTTCCATTAATCCAAAGACAATCCAGGAAGAATGCGAGGTTAAGCGGTACACG 1159
Qy 379 IleHisProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsnThr 398
Db 1160 ATTCCGCGCCAAAACTAGAATCTTCATCAACGCTGGGCTATCGGAAGGATCCCAATAC 1219
Qy 399 TrpSerGluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTyrLys 418
Db 1220 TGGGAGATCCGACACCTTCGCGCTGAGAGATTGATGAGTTTCAGGGATTTCATG 1279
Qy 419 GlyThrThrPheGluLeuValProPheGlyAlaGlyLysArgIleCysProGlyIleThr 438
Db 1280 GGAAACGATTCGAGTTTCATCCATTCGCGGCGGTTCGAAGAATCTGCCCGGTTCAT 1339
Qy 439 SerAlaIleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrpGlu 458
Db 1340 TTCGGGCTGGCAATGTTGAGATCCCATCCCTGGGCGCAACTGCTCTACCACTTCGACTG 1399
Qy 459 LeuAlaAspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyAlaLeu 478
Db 1400 TTGCCACAAGGAATGACTGATCCGACTTGGACATGACGAGACCCCGAGTCTTCTGGG 1459
Qy 479 ArgLysLysIleAspLeuLysLeuIleProIleProTyrGln 492


```
Db 1460 CCAAAAAAGAAAAATGTTGTTGGTTCCACACTCTATAA 1501
RESULT 15
; Sequence 3, Application US/08948564
; Patent No. 6121512
; GENERAL INFORMATION:
; APPLICANT: Siminezky, Balazs
; APPLICANT: Dewey, Ralph E.
; APPLICANT: Corbin, Frederick T.
; TITLE OF INVENTION: No. 6121512el Cytochrome P-450 Constructs and
; TITLE OF INVENTION: Methods of Producing Herbicide-Resistant Transgenic Plants
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Virginia C. Bennett
; STREET: PO Box 37428
; CITY: Raleigh
; STATE: No. 6121512th Carolina
; COUNTRY: USA
; ZIP: 27627
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/948,564
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Virginia C.
; REGISTRATION NUMBER: 37,092
; REFERENCE/DOCKET NUMBER: 5051-409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-854-1400
; TELEFAX: 919-854-1401
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1691 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 16..1545
US-08-948-564-3

Alignment Scores:
Pred. No.: 4,48e-98 Length: 1691
Score: 977.50 Matches: 207
Percent Similarity: 62.7% Conservative: 112
Beet Local Similarity: 40.7% Mismatches: 165
Query Match: 38.0% Indels: 25
DB: 3 Gaps: 10

US-10-759-813-2 (1-500) x US-08-948-564-3 (1-1691)
QY 1 MetGluGlnLysAsnLeuSerPheProSerIleLeuIleSerPheLeuValLeuIle 20
Db 22 ATGGAGCTTCACACACACACCCCTTCTCTATTACTTCTTACCTCCATTCTCTATT 81
QY 21 LeuValValMetArgLeuTrpLysLysGlnAsnPro----- 33
Db 82 TTCTTGGTGTCTTCAAAATTAGTTCAAGATCGGATTCACAAACCTCTCTACCTGCMAA 141
QY 34 ---ProProGlyProTrpLysPheProIleIleGlyAsnLeuProHisLeuLeuThr 52
Db 142 TTGCCCCCGAGCCAGGACACTCCTCTCATAGGGAACATA---CACCAGATTGTGGC 198
QY 53 SerAspLeuGlyHisGluArgPheArgAlaLeuAlaGlnIleTrpGlyProValMetSer 72
Db 199 TCACTGCCGGTTCATTACTTAAAAAATTTGGCAGATAAGTATGGTCCATTAAATGCAT 258

QY 73 LeuGlnIleGlyGlnValSerAlaValValIleSerSerAlaGluAlaAlaLysGluVal 92
Db 259 CTAATAACTAGGAGAGGTGTCACACATCATAGTCACTTCCCAGAAATGCGCCCAAGAGATT 318
QY 93 MetLysThrGlnAlaAspAlaPheAlaGlnArgProIleValLeuAspAlaGlnIleVal 112
Db 319 ATGAAGACACATGATCTCAACTTCTCTGATAGCCAGACTTTGTATTGTCTAGATAGTT 378
QY 113 PheTrpAsnArgLysAspValLeuPheAlaSerTrpGlyAspHisTrpArgGlnMetLys 132
Db 379 TCTTACAACGGTCTTGGCATTGCTTCTCAGTCAACATGAGACTATTGAGGCACTAAGA 438
QY 133 LysIleTrpIleLeuGluPheLeuSerAlaLysLysValGlnSerSerArgLeuIleArg 152
Db 439 AAGATATGCACAGTAGAGTTACTAACCAAGCGCGTGCAGTCTTTTCGGTCCATAAGA 498
QY 153 GluGluGluMetGluAspAlaIleThrPheLeuArgSerLysAla-----Gly 168
Db 499 GAAGAGAGGTGGCAGAACTAGTTAAAAAATAGCTGCACTGCAAGTGAAGAAGGGGG 558
QY 169 SerProValAsnIleThrLysIleLeuTrpGlyIleIleSerIleMetIleArgThr 188
Db 559 TCCATTTTAAATCTCACCAGAGCATTCTCAATGACTTTTGGGATAGCGGCAGCGG 618
QY 189 SerValGlyAsnCysLysGlnLysGluArgLeuSerValAlaAspAlaValAsnGlu 208
Db 619 GCTTTTGGTAAA---AAGAGCAGATACCAACAAGTGTTCATATCAAAATGTCATAAACAA 675
QY 209 AlaAlaThr-----SerPheGlyThrAlaAspAlaPheProThrTrpLysLeuLeuHis 226
Db 676 TTGATGCTTCTGGGAGGGTTTCTGCTGATCTCTATCTCTCTCTAGTAGAGTTTCAA 735
QY 227 TyrIleIleGlyAlaGluSerLysProArgArgLeuHisGlnGluIleAspAspLeuLeu 246
Db 736 --ATGATGGGGCGCGCGGAACTTCAAAAAGTGCATAGACTGACAGATAGGTTGTTG 792
QY 247 GluGluIleLeuAsnGluHisLys---AlaAsnLysProPheGlu-----Ala 261
Db 793 CAAGACATCATCGAGCAGACACAAAAATAGAAACAGACAGCAGGAGCGTGGAAGAGTG 852
QY 262 AspAsnLeuMetAspValLeuAsnLeuGlnLysAsnGlyAsnValProValProVal 281
Db 853 GAAGATCTAGTTGATGTTCTCTCAAGTTTCAAAAGGAATCGAA-----TTTCGCTTG 906
QY 282 ThrAsnGluSerIleLysAlaSerValLeuGlnMetPheThrAlaGlySerGluThrThr 301
Db 907 ACTGATGACAACTTAAAGCCGTCATCCAGGACATATTCTTGTGGTGGAGCGCAACATCA 966
QY 302 SerLysAlaThrGluTrpValMetAlaGluLeuMetLysAsnProThrGluLeuArgLys 321
Db 967 TCTTCTGTGTGGAAATGGGGATGTCAGAAATTTGATAAGAAACCCGAGGGTGATGGNAGAA 1026
QY 322 AlaGlnGluGluValArgGlnValPheGlyGluMetGlyLysValAspGluSerArgPhe 341
Db 1027 GCACAGCAGAGGTGAGAGAGTGTATGATACCAAGGGATATGTGGATGAGACAGAAATTG 1086
QY 342 HisAspLeuLysPhePheLysLeuValValLysGluThrLeuArgLeuHisProProVal 361
Db 1087 CACCAATTTGATATACTTAAAGTCCATCATCAAGAAACCATGAGGTATACATCCACTGTG 1146
QY 362 ---ValLeuIleProArgGluCysArgGluThrThrArgIleAspGlyTrpGluIleHis 380
Db 1147 CCATTTGTTAGTTCCTAGAGTAAGTAGAGAAAGTGCCAAATCAATGATATAGATACCC 1206
QY 381 ProAsnThrArgIleValValAsnAlaTrpAlaIleGlyArgAspProAsnThrTrpSer 400
Db 1207 TCTAAGACTAGGATCATTAATCAATGCTTGGCAATTTGGAAGGAATCCTAAGTATTGGGT 1266
QY 401 GluProGlyLysPheAsnProGluArgPheLysAspCysAlaIleAspTrpLysGlyThr 420
Db 1267 GAAACTGAGAGTGTAAACCTGAGAGGTTTCTTAATAGCTCCATTGATTTTGGGGCACA 1326
```

```
QY 421 ThrPheGluLeuValProPheGlyAlaGlyLysArgileCysProGlyIleThrSerAla 440
Db 1327 GACTTTGAATTATCCCATTTGCTGTGGAAGAGGATCTGCCCGGCATTACATTGCC 1386
QY 441 IleThrAsnLeuGluTyrValIleIleAsnLeuLeuTyrHisPheAsnTrpGluLeuAla 460
Db 1387 ATACCCCAACATTGAGTTGCCACTTGCTCAGTTACTTTACCACCTTGATTGGAAGCTTCCC 1446
QY 461 AspGlyIleThrProGlnThrLeuAspMetThrGluAlaIleGlyGlyAlaLeuArgLys 480
Db 1447 AATAAATGAAGAAATGAAGAACTGACATGACGAGTCAAAATGGAATTACTTTACGAAGA 1506
QY 481 LysIleAspLeuLysLeuIleProIle 489
Db 1507 CAAAATGACCTCTGCTTGATTTCCATT 1533
```

Search completed: May 31, 2006, 08:41:39
Job time : 329 secs

This Page Blank (uspto)